Massachusetts Institute of Technology
Department of Economics
Working Paper Series

COMPETITION AND EFFICIENCY IN
CONGESTED MARKETS

Daron Acemoglu
Asuman E. Ozdaglar

Working Paper 05-06
February 17, 2005

Room E52-251
50 Memorial Drive
Cambridge, MA 02142
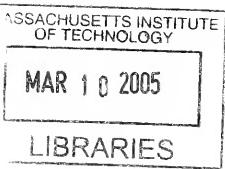
Competition and Efficiency in Congested Markets[*]

Daron Acemoglu
Department of Economics
Massachusetts Institute of Technology

Asuman E. Ozdaglar
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology

February 17, 2005

Abstract

We study the efficiency of oligopoly equilibria in congested markets. The moti-
vating examples are the allocation of network flows in a communication network or
of traffic in a transportation network. We show that increasing competition among
oligopolists can reduce efficiency, measured as the difference between users' will-
ingness to pay and delay costs. We characterize a tight bound of 5/6 on efficiency
in pure strategy equilibria. This bound is tight even when the number of routes
and oligopolists is arbitrarily large. We also study the efficiency properties of
mixed strategy equilibria.

1

1   Introduction

We analyze price competition in the presence of congestion costs. Consider the following environment: one unit of traffic can use one of I alternative routes. More traffic on a particular route causes delays, exerting a negative (congestion) externality on existing traffic.[1] Congestion costs are captured by a route-specific non-decreasing convex latency function, $l_i(\cdot)$. Profit-maximizing oligopolists set prices (tolls) for travel on each route denoted by p_i. We analyze subgame perfect Nash equilibria of this environment, where for each price vector, p, all traffic chooses the path that has minimum (toll plus delay) cost, $l_i + p_i$, and oligopolists choose prices to maximize profits.

The environment we analyze is of practical importance for a number of settings. These include transportation and communication networks, where additional use of a route (path) generates greater congestion for all users, and markets in which there are "snob" effects, so that goods consumed by fewer other consumers are more valuable (see for example, [38]). The key feature of these environments is the negative congestion externality that users exert on others. This externality has been well-recognized since the work by Pigou [27] in economics, by Samuelson [33], Wardrop [41], Beckmann et al. [4] in transportation networks, and by Orda et. al. [23], Korilis et. al. [18], Kelly [17], Low [20] in communication networks. More recently, there has been a growing literature that focuses on quantification of efficiency loss (referred to as the *price of anarchy*) that results from externalities and strategic behavior in different classes of problems; selfish routing (Koutsoupias and Papadimitriou [19], Roughgarden and Tardos [31], Correa, Schulz, and Stier-Moses [8], Perakis [26], and Friedman [12]), resource allocation by market mechanisms (Johari and Tsitsiklis [16], Sanghavi and Hajek [32]), and network design (Anshelevich et.al. [2]). Nevertheless, the game-theoretic interactions between (multiple) service providers and users, or the effects of competition among the providers on the efficiency loss has not been considered. This is an important area for analysis since in most applications, (competing) profit-maximizing entities charge prices for use. Moreover, we will show that the nature of the analysis changes significantly in the presence of price competition.

We provide a general framework for the analysis of price competition among providers in a congested (and potentially capacitated) network, study existence of pure strategy and mixed strategy equilibria, and characterize and quantify the efficiency properties of equilibria. There are four sets of major results from our analysis.

First, though the equilibrium of traffic assignment without prices can be highly inefficient (e.g., [27], [31], [8]), price-setting by a monopolist internalizes the negative externality and achieves efficiency.

Second, increasing competition can increase inefficiency. In fact, changing the market structure from monopoly to duopoly almost always increases inefficiency. This result contrasts with most existing results in the economics literature where greater competition tends to improve the allocation of resources (e.g. see Tirole [36]). The intuition for this result is driven by the presence of congestion and is illustrated by the example we discuss below.[2]

---

[1] An externality arises when the actions of the player in a game affects the payoff of other players.

[2] Because users are homogeneous and have a constant reservation utility in our model, in the absence

1

Third, we provide a tight bound on the extent of inefficiency in the presence of pricing, which applies irrespective of the number of routes, I. We show that social surplus (defined as the difference between users' willingness to pay and the delay cost) in any pure strategy oligopoly equilibrium is always greater than 5/6 of the maximum social surplus. Simple examples reach this 5/6 bound. Interestingly, this bound is obtained even when the number of routes, I, is arbitrarily large.

Fourth, pure strategy equilibria may fail to exist. This is not surprising in view of the fact that what we have here is a version of a Bertrand-Edgeworth game where pure strategy equilibria do not exist in the presence of convex costs of production or capacity constraints (e.g., Edgeworth [11], Shubik [35], Benassy [6], Vives [40]). However, in our oligopoly environment we show that when latency functions are linear, a pure strategy equilibrium always exists, essentially because some amount of congestion externalities remove the payoff discontinuities inherent in the Bertrand-Edgeworth game. Non-existence becomes an issue in our environment when latency functions are highly convex. In this case, we prove that mixed strategy equilibria always exist. We also show that mixed strategy equilibria can lead to arbitrarily inefficient worst-case realizations; in particular, social surplus can become arbitrarily small relative to the maximum social surplus.

The following example illustrates some of these results.

Example 1 Figure 1 shows a situation similar to the one first analyzed by Pigou [27] to highlight the inefficiency due to congestion externalities. One unit of traffic will travel from origin A to destination B, using either route 1 or route 2. The latency functions are given by

$$l_1(x) = \frac{x^2}{3}, \qquad l_2(x) = \frac{2}{3}x.$$

It is straightforward to see that the efficient allocation [i.e., one that minimizes the total delay cost $\sum_i l_i(x_i)x_i$] is $x_1^S = 2/3$ and $x_2^S = 1/3$, while the (Wardrop) equilibrium allocation that equates delay on the two paths is $x_1^{WE} \approx .73 > x_1^S$ and $x_2^{WE} \approx .27 < x_2^S$. The source of the inefficiency is that each unit of traffic does not internalize the greater increase in delay from travel on route 1, so there is too much use of this route relative to the efficient allocation.

Now consider a monopolist controlling both routes and setting prices for travel to maximize its profits. We will show in greater detail below that in this case, the monopolist will set a price including a markup, $x_i l_i'$ (when l_i is differentiable), which exactly internalizes the congestion externality. In other words, this markup is equivalent to the Pigovian tax that a social planner would set in order to induce decentralized traffic to choose the efficient allocation. Consequently, in this simple example, monopoly prices will be $p_1^{ME} = (2/3)^3 + k$ and $p_2^{ME} = (2/3^2) + k$, for some constant k. The resulting traffic in the Wardrop equilibrium will be identical to the efficient allocation, i.e., $x_1^{ME} = 2/3$ and $x_2^{ME} = 1/3$.

Next consider a duopoly situation, where each route is controlled by a different profit-maximizing provider. In this case, it can be shown that equilibrium prices will take the

---

of congestion externalities, all market structures would achieve efficiency, and a change from monopoly to duopoly, for example, would have no efficiency consequence.

$$l_1(x) = x^2/3$$

1 unit of traffic

$$l_2(x) = (2/3)x$$
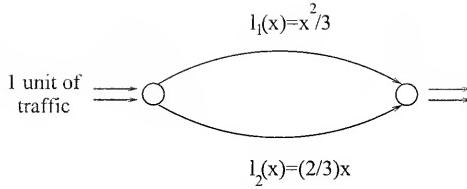
Figure 1: A two link network with congestion-dependant latency functions.

form $p_i^{OE} = x_i(l_1' + l_2')$ [see Eq. (20) in Section 4], or more specifically, $p_1^{OE} \approx 0.61$ and $p_2^{OE} \approx 0.44$. The resulting equilibrium traffic is $x_1^{OE} \approx .58 < x_1^S$ and $x_2^{OE} \approx .42 > x_2^S$, which also differs from the efficient allocation. We will show that this is generally the case in the oligopoly equilibrium. Interestingly, while in the Wardrop equilibrium without prices, there was too much traffic on route 1, now there is too little traffic because of its greater markup. It is also noteworthy that although the duopoly equilibrium is inefficient relative to the monopoly equilibrium, in the monopoly equilibrium k is chosen such that all of the consumer surplus is captured by the monopolist, while in the oligopoly equilibrium users may have positive consumer surplus.[3]

The intuition for the inefficiency of duopoly relative to monopoly can be obtained as follows. There is now a new source of (differential) monopoly power for each duopolist, which they exploit by distorting the pattern of traffic: when provider 1, controlling route 1, charges a higher price, it realizes that this will push some traffic from route 1 to route 2, raising congestion on route 2. But this makes the traffic using route 1 become more "locked-in," because their outside option, travel on the route 2, has become worse.[4] As a result, the optimal price that each duopolist charges will include an additional markup over the Pigovian markup. These are $x_1 l_2'$ for route 1 and $x_2 l_1'$ for route 2. Since these two markups are generally different, they will distort the pattern of traffic away from the efficient allocation. Naturally, however, prices are typically lower with duopoly, so even though social surplus declines, users will be better off than in monopoly (i.e., they will command a positive consumer surplus).

Although there is a large literature on models of congestion both in transportation and communication networks, very few studies have investigated the implications of having the "property rights" over routes assigned to profit-maximizing providers. In [3], Basar and Srikant analyze monopoly pricing under specific assumptions on the utility and latency functions. He and Walrand [15] study competition and cooperation among internet service providers under specific demand models. Issues of efficient allocation of

---

[3]Consumer surplus is defined as the difference between users' willingness to pay and effective costs, $p_i + l_i(x_i)$, and is thus different from social surplus (which is the difference between users' willingness to pay and latency cost, $l_i(x_i)$, thus also takes into account producer surplus/profits). See Mas-Colell, Winston, and Green [21].

[4]Using economics terminology, we could also say that the demand for route 1 becomes more "inelastic". Since this term has a different meaning in the communication networks literature (see [34]), we do not use it here.

flows or traffic across routes do not arise in these papers. Our previous work [1] studies the monopoly problem and contains the efficiency of the monopoly result (with inelastic traffic and more restrictive assumptions on latencies), but none of the other results here. To minimize overlap, we discuss the monopoly problem only briefly in this paper.

In the rest of the paper, we use the terminology of a (communication) network, though all of the analysis applies to resource allocation in transportation networks, electricity markets, and other economic applications. Section 2 describes the basic environment. Section 3 briefly characterizes the monopoly equilibrium and establishes its efficiency. Section 4 defines and characterizes the oligopoly equilibria with competing profit-maximizing providers. Section 5 contains the main results and characterizes the efficiency properties of the oligopoly equilibrium and provide bounds on efficiency. Section 6 contains concluding comments.

Regarding notation, all vectors are viewed as column vectors, and inequalities are to be interpreted componentwise. We denote by $\mathbb{R}_+^I$ the set of nonnegative I-dimensional vectors. Let C_i be a closed subset of $[0, \infty)$ and let $f : C_i \mapsto \mathbb{R}$ be a convex function. We use $\partial f(x)$ to denote the set of subgradients of f at x, and $f^-(x)$ and $f^+(x)$ to denote the left and right derivatives of f at x.

2 Model

We consider a network with I parallel links. Let $\mathcal{I} = \{1, \ldots, I\}$ denote the set of links. Let x_i denote the total flow on link i, and $x = [x_1, \ldots, x_I]$ denote the vector of link flows. Each link in the network has a flow-dependent latency function $l_i(x_i)$, which measures the travel time (or delay) as a function of the total flow on link i. We denote the price per unit flow (bandwidth) of link i by p_i. Let $p = [p_1, \ldots, p_I]$ denote the vector of prices.

We are interested in the problem of routing d units of flow across the I links. We assume that this is the aggregate flow of many "small" users and thus adopt the Wardrop's principle (see [41]) in characterizing the flow distribution in the network; i.e., the flows are routed along paths with minimum effective cost, defined as the sum of the latency at the given flow and the price of that path (see the definition below). Wardrop's principle is used extensively in modelling traffic behavior in transportation networks ([4], [9], [25]) and communication networks ([31], [8]). We also assume that the users have a *reservation utility* R and decide not to send their flow if the effective cost exceeds the reservation utility. This implies that user preferences can be represented by the piecewise linear aggregate utility function $u(\cdot)$ depicted in Figure 2.[5]

Definition 1 For a given price vector $p \geq 0$,[6] a vector $x^{WE} \in \mathbb{R}_+^I$ is a *Wardrop*

---

[5]This simplifying assumption implies that all users are "homogeneous" in the sense that they have the same reservation utility, R. We discuss potential issues in extending this work to users with elastic and heterogeneous requirements in the concluding section.

[6]Since the reservation utility of users is equal to R, we can also restrict attention to $p_i \leq R$ for all i. Throughout the paper, we use $p \geq 0$ and $p \in [0, R]^I$ interchangeably.
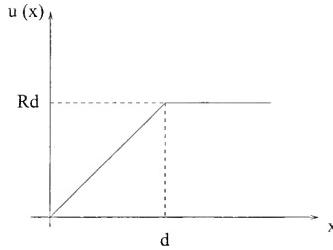
Figure 2: Aggregate utility function.

equilibrium (WE) if

$$x^{WE} \in \arg \max_{\substack{x \geq 0 \\ \sum_{i=1}^{I} x_i \leq d}} \left\{ \sum_{i=1}^{I} (R - l_i(x_i^{WE}) - p_i)x_i \right\}. \tag{1}$$

We denote the set of WE at a given p by $W(p)$.

We adopt the following assumption on the latency functions throughout the paper.

Assumption 1 For each $i \in \mathcal{I}$, the latency function $l_i : [0, \infty) \mapsto [0, \infty]$ is proper closed convex, nondecreasing, and satisfies $l_i(0) = 0$.

Hence, we allow the latency functions to be extended real-valued (see [7]). Let $C_i = \{x \in [0, \infty) \mid l_i(x) < \infty\}$ denote the effective domain of l_i. By Assumption 1, C_i is a closed interval of the form $[0, b]$ or $[0, \infty)$. Let $b_{C_i} = \sup_{x \in C_i} x$. Without loss of generality, we can add the constraint $x_i \in C_i$ in Eq. (1). Using the optimality conditions for problem (1), we see that a vector $x^{WE} \in \mathbb{R}_+^I$ is a WE if and only if $\sum_{i=1}^{I} x_i^{WE} \leq d$ and there exists some $\lambda \geq 0$ such that $\lambda(\sum_{i=1}^{I} x_i^{WE} - d) = 0$ and for all i,

$$\begin{aligned} R - l_i(x_i^{WE}) - p_i \quad &\leq \lambda \quad \text{if } x_i^{WE} = 0, \\ &= \lambda \quad \text{if } 0 < x_i^{WE} < b_{C_i}, \\ &\geq \lambda \quad \text{if } x_i^{WE} = b_{C_i}. \end{aligned} \tag{2}$$

When the latency functions are real-valued [i.e., $C_i = [0, \infty)$], we obtain the following characterization of a WE, which is often used as the definition of a WE in the literature. This lemma states that in the WE, *the effective costs*, defined as $l_i(x_i^{WE}) + p_i$, are equalized on all links with positive flows (proof omitted).

Lemma 1 Let Assumption 1 hold, and assume further that $C_i = [0, \infty)$ for all $i \in \mathcal{I}$. Then a nonnegative vector $x^* \in W(p)$ if and only if

$$l_i(x_i^*) + p_i = \min_j \{l_j(x_j^*) + p_j\}, \qquad \forall \, i \text{ with } x_i^* > 0, \tag{3}$$

5

$$l_i(x_i^*) + p_i \;\; \leq \;\; R, \qquad \forall\, i \text{ with } x_i^* > 0,$$

$$\sum_{i=1}^{I} x_i^* \;\; \leq \;\; d,$$

with $\sum_{i=1}^{I} x_i^* = d$ if $\min_j \{l_j(x_j) + p_j\} < R$.

Example 2 below shows that condition (3) in this lemma may not hold when the latency functions are not real-valued. Next we establish the existence of a WE.

Proposition 1 **(Existence and Continuity)** Let Assumption 1 hold. For any price vector $p \geq 0$, the set of WE, $W(p)$, is nonempty. Moreover, the correspondence $W : \mathbb{R}_+^I \rightrightarrows \mathbb{R}_+^I$ is upper semicontinuous.

Proof. Given any $p \geq 0$, consider the following optimization problem

$$\text{maximize}_{x \geq 0} \quad \sum_{i=1}^{I} \left((R - p_i)x_i - \int_0^{x_i} l_i(z)dz \right) \tag{4}$$

$$\text{subject to} \quad \sum_{i=1}^{I} x_i \leq d.$$

$$x_i \in C_i, \qquad \forall\, i.$$

In view of Assumption (1) (i.e., l_i is nondecreasing for all i), it can be shown that the objective function of problem (4) is convex over the constraint set, which is nonempty (since $0 \in C_i$) and convex. Moreover, the first order optimality conditions of problem (4), which are also sufficient conditions for optimality, are identical to the WE optimality conditions [cf. Eq. (2)]. Hence a flow vector $x^{WE} \in W(p)$ if and only if it is an optimal solution of problem (4). Since the objective function of problem (4) is continuous and the constraint set is compact, this problem has an optimal solution, showing that $W(p)$ is nonempty. The fact that W is an upper semicontinuous correspondence at every p follows by using the Theorem of the Maximum (see Berge [5], chapter 6) for problem (4). **Q.E.D.**

WE flows satisfy intuitive monotonicity properties given in the following proposition. The proof follows from the optimality conditions [cf. Eq. (2)] and is omitted (see [1]).

Proposition 2 **(Monotonicity)** Let Assumption 1 hold. For a given $p \geq 0$, let $p_{-j} = [p_i]_{i \neq j}$.

(a) For some $\bar{p} \leq p$, let $\bar{x} \in W(\bar{p})$ and $x \in W(p)$. Then, $\sum_{i=1}^{I} \bar{x}_i \geq \sum_{i=1}^{I} x_i$.

(b) For some $\bar{p}_j < p_j$, let $\bar{x} \in W(\bar{p}_j, p_{-j})$ and $x \in W(p_j, p_{-j})$. Then $\bar{x}_j \geq x_j$ and $\bar{x}_i \leq x_i$, for all $i \neq j$.

(c) For some $\widetilde{\mathcal{I}} \subset \mathcal{I}$, suppose that $\bar{p}_j < p_j$ for all $j \in \widetilde{\mathcal{I}}$ and $\bar{p}_j = p_j$ for all $j \notin \widetilde{\mathcal{I}}$, and let $\bar{x} \in W(\bar{p})$ and $x \in W(p)$. Then $\sum_{j \in \widetilde{\mathcal{I}}} x_j(\bar{p}) \geq \sum_{j \in \widetilde{\mathcal{I}}} x_j(p)$.

6

For a given price vector p, the WE need not be unique in general. The following example illustrates some properties of the WE.

Example 2 Consider a two link network. Let the total flow be $d = 1$ and the reservation utility be $R = 1$. Assume that the latency functions are given by

$$l_1(x) = l_2(x) = \begin{cases} 0 & \text{if } 0 \le x \le \frac{2}{3} \\ \infty & \text{otherwise.} \end{cases}$$

At the price vector $(p_1, p_2) = (1, 1)$, the set of WE, $W(p)$, is given by the set of all vectors (x_1, x_2) with $0 \le x_i \le 2/3$ and $\sum_i x_i \le 1$. At any price vector (p_1, p_2) with $p_1 > p_2 = 1$, $W(p)$ is given by all $(0, x_2)$ with $0 \le x_2 \le 2/3$.

This example also illustrates that Lemma 1 need not hold when latency functions are not real-valued. Consider, for instance, the price vector $(p_1, p_2) = (1 - \epsilon, 1 - a\epsilon)$ for some scalar $a > 1$. In this case, the unique WE is $(x_1, x_2) = (1/3, 2/3)$, and clearly effective costs on the two routes are not equalized despite the fact that they both have positive flows. This arises because the path with the lower effective cost is capacity constrained, so no more traffic can use that path.

Under further restrictions on the l_i, we obtain the following standard result in the literature which assumes strictly increasing latency functions.

Proposition 3 (Uniqueness) Let Assumption 1 hold. Assume further that l_i is strictly increasing over C_i. For any price vector $p \ge 0$, the set of WE, $W(p)$, is a singleton. Moreover, the function $W : \mathbb{R}_+^I \mapsto \mathbb{R}_+^I$ is continuous.

Proof. Under the given assumptions, for any $p \ge 0$, the objective function of problem (4) is strictly convex, and therefore has a unique optimal solution. This shows the uniqueness of the WE at a given p. Since the correspondence W is upper semicontinuous from Proposition 1 and single-valued, it is continuous. **Q.E.D.**

In general, under the Assumption that $l_i(0) = 0$ for all $i \in \mathcal{I}$, we have the following result, which is essential in our analysis with nonunique WE flows.

Lemma 2 Let Assumption 1 hold. For a given $p \ge 0$, define the set

$$\bar{\mathcal{I}} = \{i \in \mathcal{I} \mid \exists \, x, \, \bar{x} \in W(p) \text{ with } x_i \ne \bar{x}_i\}. \tag{5}$$

Then

$$\begin{aligned} l_i(x_i) &= 0, & \forall \, i \in \bar{\mathcal{I}}, \, \forall \, x \in W(p), \\ p_i &= p_j, & \forall \, i, \, j \in \bar{\mathcal{I}}. \end{aligned}$$

Proof. Consider some $i \in \bar{\mathcal{I}}$ and $x \in W(p)$. Since $i \in \bar{\mathcal{I}}$, there exists some $\hat{x} \in W(p)$ such that $x_i \neq \hat{x}_i$. Assume without loss of generality that $x_i > \hat{x}_i$. There are two cases to consider:

(a) If $x_k \geq \hat{x}_k$ for all $k \neq i$, then $\sum_{j \in \mathcal{I}} x_j > \sum_{j \in \mathcal{I}} \hat{x}_j$, which implies that the WE optimality conditions [cf. Eq. (2)] for $\hat{x}$ hold with $\hat{\lambda} = 0$. By Eq. (2) and $x_i > \hat{x}_i$, we have

$$l_i(x_i) + p_i \leq R,$$

$$l_i(\hat{x}_i) + p_i \geq R,$$

which together imply that $l_i(x_i) = l_i(\hat{x}_i)$. By Assumption 1 (i.e., l_i is convex and $l_i(0) = 0$), it follows that $l_i(x_i) = 0$.

(b) If $x_k < \hat{x}_k$ for some k, by the WE optimality conditions, we obtain

$$l_i(x_i) + p_i \leq l_k(x_k) + p_k,$$

$$l_i(\hat{x}_i) + \hat{p}_i \geq l_k(\hat{x}_k) + p_k.$$

Combining the above with $x_i > \hat{x}_i$ and $x_k < \hat{x}_k$, we see that $l_i(x_i) = l_i(\hat{x}_i)$, and $l_k(x_k) = l_k(\hat{x}_k)$. By Assumption 1, this shows that $l_i(x_i) = 0$ (and also that $p_i = p_k$).

Next consider some $i, j \in \bar{\mathcal{I}}$. We will show that $p_i = p_j$. Since $i \in \bar{\mathcal{I}}$, there exist $x, \hat{x} \in W(p)$ such that $x_i > \hat{x}_i$. There are three cases to consider:

- $x_j < \hat{x}_j$. Then a similar argument to part (b) above shows that $p_i = p_j$.

- $x_j > \hat{x}_j$. If $x_k \geq \hat{x}_k$ for all $k \neq i, j$, then $\sum_m \hat{x}_m < d$, implying that the WE optimality conditions hold with $\hat{\lambda} = 0$. Therefore, we have

$$l_i(x_i) + p_i \leq R,$$

$$l_j(\hat{x}_j) + p_j \geq R,$$

which together with $l_i(x_i) = l_j(\hat{x}_j) = 0$ imply that $p_i = p_j$.

- $x_j = \hat{x}_j$. Since $j \in \bar{\mathcal{I}}$, there exists some $\bar{x} \in W(p)$ such that $x_j \neq \bar{x}_j$. Repeating the above two steps with $\bar{x}_j$ instead of $\hat{x}_j$ yields the desired result.

Q.E.D.

We next define the social problem and the social optimum, which is the routing (flow allocation) that would be chosen by a planner that has full information and full control over the network.

Definition 2 A flow vector x^S is a *social optimum* if it is an optimal solution of the *social problem*

$$\text{maximize}_{x \geq 0} \quad \sum_{i=1}^{I} \Big(R - l_i(x_i) \Big) x_i \tag{6}$$

$$\text{subject to} \quad \sum_{i=1}^{I} x_i \leq d.$$

In view of Assumption 1, the social problem has a continuous objective function and a compact constraint set, guaranteeing the existence of a social optimum, x^S. Moreover, using the optimality conditions for a convex program (see [7], Section 4.7), we see that a vector $x^S \in \mathbb{R}_+^I$ is a social optimum if and only if $\sum_{i=1}^{I} x_i^S \leq d$ and there exists a subgradient $g_{l_i} \in \partial l_i(x_i^S)$ for each i, and a $\lambda^S \geq 0$ such that $\lambda^S (\sum_{i=1}^{I} x_i^S - d) = 0$ and for each i,

$$
\begin{aligned}
R - l_i(x_i^S) - x_i^S g_{l_i} \quad &\leq \lambda^S \quad \text{if } x_i^S = 0, \\
&= \lambda^S \quad \text{if } 0 < x_i^S < b_{C_i}, \\
&\geq \lambda^S \quad \text{if } x_i^S = b_{C_i}.
\end{aligned}
\tag{7}
$$

For future reference, for a given vector $x \in \mathbb{R}_+^I$, we define the value of the objective function in the social problem,

$$\mathbb{S}(x) = \sum_{i=1}^{I} \left(R - l_i(x_i) \right) x_i, \tag{8}$$

as the *social surplus*, i.e., the difference between users' willingness to pay and the total latency.

3 Monopoly Equilibrium and Efficiency

In this section, we assume that a monopolist service provider owns the I links and charges a price of p_i per unit bandwidth on link i. We considered a related problem in Acemoglu and Ozdaglar [1] for atomic users with inelastic traffic (i.e., the utility function of each of a finite set of users is a step function), and with increasing, real-valued and differentiable latency functions. Here we show that similar results hold for the more general latency functions and the demand model considered in Section 2.

The monopolist sets the prices to maximize his profit given by

$$\Pi(p, x) = \sum_{i=1}^{I} p_i x_i,$$

where $x \in W(p)$. This defines a two-stage dynamic *pricing-congestion game*, where the monopolist sets prices anticipating the demand of users, and given the prices (i.e., in each subgame), users choose their flow vectors according to the WE.

Definition 3 A vector $(p^{ME}, x^{ME}) \geq 0$ is a *Monopoly Equilibrium* (ME) if $x^{ME} \in W(p^{ME})$ and

$$\Pi(p^{ME}, x^{ME}) \geq \Pi(p, x), \qquad \forall\, p,\ \forall\, x \in W(p).$$

Our definition of the ME is stronger than the standard subgame perfect Nash equilibrium concept for dynamic games. With a slight abuse of terminology, let us associate a subgame perfect Nash equilibrium with the on-the-equilibrium-path actions of the two-stage game.

Definition 4 A vector $(p^*, x^*) \geq 0$ is a subgame perfect equilibrium (SPE) of the pricing-congestion game if $x^* \in W(p^*)$ and for all $p \geq 0$, there exists $x \in W(p)$ such that

$$\Pi(p^*, x^*) \geq \Pi(p, x).$$

The following proposition shows that under Assumption 1, the two solution concepts coincide. Since the proof is not relevant for the rest of the argument, we provide it in Appendix A.

Proposition 4 Let Assumption 1 hold. A vector (p^{ME}, xME) is an ME if and only if it is an SPE of the pricing-congestion game.

Since an ME (p^*, x^*) is an optimal solution of the optimization problem

$$\text{maximize}_{p \geq 0,\ x \geq 0} \quad \sum_{i=1}^{I} p_i x_i \tag{9}$$
$$\text{subject to} \quad x \in W(p),$$

it is slightly easier to work with than an SPE. Therefore, we use ME as the solution concept in this paper.

The preceding problem has an optimal solution, which establishes the existence of an ME. In the next proposition, we show that the flow allocation at an ME and the social optimum are the same.

Proposition 5 Let Assumption 1 hold. A vector x is the flow vector at an ME if and only if it is a social optimum.

To prove this result, we first establish the following lemma.

Lemma 3 Let Assumption 1 hold. Let (p, x) be an ME. Then for all i with $x_i > 0$, we have

$$p_i = R - l_i(x_i).$$

Proof. Define $\bar{\mathcal{I}} = \{i \in \mathcal{I} \mid x_i > 0\}$. By the definition of a WE, for all $i \in \bar{\mathcal{I}}$, we have $p_i + l_i(x_i) \le R$. We first show that

$$l_i(x_i) + p_i = l_j(x_j) + p_j, \qquad \forall\, i, j \in \bar{\mathcal{I}}.$$

Suppose $l_i(x_i) + p_i < l_j(x_j) + p_j$ for some $i, j \in \bar{\mathcal{I}}$. Using the optimality conditions for a WE [cf. Eq. (2)], we see that $x \in W(p_i + \epsilon, p_{-i})$ for sufficiently small ϵ, contradicting the fact that (p, x) is an ME. Assume next that $p_i + l_i(x_i) < R$ for some $i \in \bar{\mathcal{I}}$. Consider the price vector $\tilde{p} = p + \epsilon \sum_{i \in \bar{\mathcal{I}}} e_i$, where e_i is the ith unit vector and ϵ is such that

$$p_i + \epsilon + l_i(x_i) < R, \qquad \forall\, i \in \bar{\mathcal{I}}.$$

Hence, x is a WE at price $\tilde{p}$, and therefore the vector $(\tilde{p}, x)$ is feasible for problem (9) and has a strictly higher objective function value, contradicting the fact that (p, x) is an ME. **Q.E.D.**

Proof of Proposition 5: In view of Lemma 3, we can rewrite problem (9) as

$$\text{maximize}_{x \ge 0} \quad \sum_{i=1}^{I} \Big(R - l_i(x_i) \Big) x_i$$

$$\text{subject to} \quad \sum_{i=1}^{I} x_i \le d.$$

This problem is identical to the social problem [cf. problem (6)], completing the proof. **Q.E.D.**

In addition to the social surplus defined above, it is also useful to define the *consumer surplus*, as the difference between users' willingness to pay and effective cost, i.e., $\sum_{i=1}^{I} (R - l_i(x_i) - p_i)x_i$ (See Mas-Colell, Winston and Green, [21]). By Proposition 5 and Lemma 3, it is clear that even though the ME achieves the social optimum, all of the surplus is captured by the monopolist, and users are just indifferent between sending their information or not (i.e., receive no consumer surplus).

Our major motivation for the study of oligopolistic settings is that they provide a better approximation to reality, where there is typically competition among service providers. A secondary motivation is to see whether an oligopoly equilibrium will achieve an efficient allocation like the ME, while also transferring some or all of the surplus to the consumers.

4 Oligopoly Equilibrium

We suppose that there are S service providers, denote the set of service providers by $\mathcal{S}$, and assume that each service provider $s \in \mathcal{S}$ owns a different subset $\mathcal{I}_s$ of the links. Service provider s charges a price p_i per unit bandwidth on link $i \in \mathcal{I}_s$. Given the vector

of prices of links owned by other service providers, $p_{-s} = [p_i]_{i \notin \mathcal{I}_s}$, the profit of service provider s is

$$\Pi_s(p_s, p_{-s}, x) = \sum_{i \in \mathcal{I}_s} p_i x_i,$$

for $x \in W(p_s, p_{-s})$, where $p_s = [p_i]_{i \in \mathcal{I}_s}$.

The objective of each service provider, like the monopolist in the previous section, is to maximize profits. Because their profits depend on the prices set by other service providers, each service provider forms conjectures about the actions of other service providers, as well as the behavior of users, which, we assume, they do according to the notion of (subgame perfect) Nash equilibrium. We refer to the game among service providers as the *price competition game*.

Definition 5 A vector $(p^{OE}, x^{OE}) \geq 0$ is a (pure strategy) *Oligopoly Equilibrium* (OE) if $x^{OE} \in W\left(p_s^{OE}, p_{-s}^{OE}\right)$ and for all $s \in \mathcal{S}$,

$$\Pi_s(p_s^{OE}, p_{-s}^{OE}, x^{OE}) \geq \Pi_s(p_s, p_{-s}^{OE}, x), \qquad \forall\, p_s \geq 0,\ \forall\, x \in W(p_s, p_{-s}^{OE}). \qquad (10)$$

We refer to p^{OE} as the *OE price*.

As for the monopoly case, there is a close relation between a pure strategy OE and a pure strategy subgame perfect equilibrium. Again associating the subgame perfect equilibrium with the on-the-equilibrium-path actions, we have the following standard definition.

Definition 6 A vector $(p^*, x^*) \geq 0$ is a subgame perfect equilibrium (SPE) of the price competition game if $x^* \in W(p^*)$ and there exists a function $x(p) \in W(p)$ such that for all $s \in \mathcal{S}$,

$$\Pi_s(p_s^*, p_{-s}^*, x^*) \geq \Pi_s(p_s, p_{-s}^*, x\left(p_s, p_{-s}^*\right)) \qquad \forall\, p_s \geq 0. \qquad (11)$$

The following proposition generalizes Proposition 4 and enables us to work with the OE definition, which is more convenient for the subsequent analysis. The proof parallels that of Proposition 4 and is omitted.

Proposition 6 Let Assumption 1 hold. A vector (p^{OE}, x^{OE}) is an OE if and only if it is an SPE of the price competition game.

The price competition game is neither concave nor supermodular. Therefore, classical arguments that are used to show the existence of a pure strategy equilibrium do not hold (see [13], [37]). In the next proposition, we show that for linear latency functions, there exists a pure strategy OE.

Proposition 7 Let Assumption 1 hold, and assume further that the latency functions are linear. Then the price competition game has a pure strategy OE.

Proof. For all $i \in \mathcal{I}$, let $l_i(x) = a_i x$. Define the set

$$\mathcal{I}_0 = \{i \in \mathcal{I} \mid a_i = 0\}.$$

Let I_0 denote the cardinality of set $\mathcal{I}_0$. There are two cases to consider:

- $I_0 \geq 2$: Assume that there exist i, $j \in \mathcal{I}_0$ such that $i \in \mathcal{I}_s$ and $j \in \mathcal{I}_{s'}$ for some $s \neq s' \in \mathcal{S}$. Then it can be seen that a vector (p^{OE}, x^{OE}) with $p_i^{OE} = 0$ for all $i \in \mathcal{I}_0$ and $x^{OE} \in W(p^{OE})$ is an OE. Assume next that for all $i \in \mathcal{I}_0$, we have $i \in \mathcal{I}_s$ for some $s \in \mathcal{S}$. Then, we can assume without loss of generality that provider s owns a single link i' with $a_{i'} = 0$ and consider the case $I_0 = 1$.

- $I_0 \leq 1$: Let $B_s(p_{-s}^{OE})$ be the set of p_s^{OE} such that

$$(p_s^{OE}, x^{OE}) \in \arg \max_{\substack{p_s \geq 0 \\ x \in W(p_s, p_{-s}^{OE})}} \sum_{i \in \mathcal{I}_s} p_i x_i. \tag{12}$$

Let $B(p^{OE}) = [B_s(p_{-s}^{OE})]_{s \in \mathcal{S}}$. In view of the linearity of the latency functions, it follows that $B(p^{OE})$ is an upper semicontinuous and convex-valued correspondence. Hence, we can use Kakutani's fixed point theorem to assert the existence of a p^{OE} such that $B(p^{OE}) = p^{OE}$ (see Berge [5]). To complete the proof, it remains to be shown that there exists $x^{OE} \in W(p^{OE})$ such that Eq. (10) holds.

If $\mathcal{I}_0 = \emptyset$, we have by Proposition 3 that $W(p^{OE})$ is a singleton, and therefore Eq. (10) holds and $(p^{OE}, W(p^{OE}))$ is an OE.

Assume finally that exactly one of the a_i's (without loss of generality a_1) is equal to 0. We show that for all $\bar{x}$, $\tilde{x} \in W(p^{OE})$, we have $\bar{x}_i = \tilde{x}_i$, for all $i \neq 1$. Let $EC(x, p^{OE}) = \min_j \{l_j(x_j) + p_j^{OE}\}$. If at least one of

$$EC(\tilde{x}, p^{OE}) < R, \qquad \text{or} \qquad EC(\bar{x}, p^{OE}) < R$$

holds, then one can show that $\sum_{i=1}^{I} \tilde{x}_i = \sum_{i=1}^{I} \bar{x}_i = d$. Substituting $x_1 = d - \sum_{i \in \mathcal{I}, \, i \neq 1} x_i$ in problem (4), we see that the objective function of problem (4) is strictly convex in $x_{-1} = [x_i]_{i \neq 1}$, thus showing that $\tilde{x} = \bar{x}$. If both $EC(\tilde{x}, p^{OE}) = R$ and $EC(\bar{x}, p^{OE}) = R$, then $\bar{x}_i = \tilde{x}_i = l_i^{-1}(R - p_i^{OE})$ for all $i \neq 1$, establishing our claim.

For some $x \in W(p^{OE})$, consider the vector $x^{OE} = (d - \sum_{i \neq 1} x_i, x_{-1})$. Since x_{-1} is uniquely defined and x_1 is chosen such that the provider that owns link 1 has no incentive to deviate, it follows that (p^{OE}, x^{OE}) is an OE.

Q.E.D.

The existence result cannot be generalized to piecewise linear latency functions or to latency functions which are linear over their effective domain, as illustrated in the following example.

Example 3 Consider a two link network. Let the total flow be $d = 1$. Assume that the latency functions are given by

$$l_1(x) = 0, \qquad l_2(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \delta \\ \frac{x-\delta}{\epsilon} & x \geq \delta, \end{cases}$$

for some $\epsilon > 0$ and $\delta > 1/2$, with the convention that when $\epsilon = 0$, $l_2(x) = \infty$ for $x > \delta$. We first show that there exists no pure strategy oligopoly equilibrium for small ϵ (i.e., there exists no pure strategy subgame perfect equilibrium). The following list considers all candidate oligopoly price equilibria (p_1, p_2) and profitable unilateral deviations for ϵ sufficiently small, thus establishing the nonexistence of an OE:

1. $p_1 = p_2 = 0$: A small increase in the price of provider 1 will generate positive profits, thus provider 1 has an incentive to deviate.

2. $p_1 = p_2 > 0$: Let x be the flow allocation at the OE. If $x_1 = 1$, then provider 2 has an incentive to decrease its price. If $x_1 < 1$, then provider 1 has an incentive to decrease its price.

3. $0 \leq p_1 < p_2$: Player 1 has an incentive to increase its price since its flow allocation remains the same.

4. $0 \leq p_2 < p_1$: For ϵ sufficiently small, the profit function of player 2, given p_1, is strictly increasing as a function of p_2, showing that provider 2 has an incentive to increase its price.

We next show that a mixed strategy OE always exists. We define a mixed strategy OE as a mixed strategy subgame perfect equilibrium of the price competition game (see Dasgupta and Maskin, [10]). Let $\mathcal{B}^n$ be the space of all (Borel) probability measures on $[0, R]^n$. Let I_s denote the cardinality of $\mathcal{I}_s$, i.e., the number of links controlled by service provider s. Let $\mu_s \in \mathcal{B}^{I_s}$ be a probability measure, and denote the vector of these probability measures by μ and the vector of these probability measures excluding s by μ_{-s}.

Definition 7 $(\mu^*, x^*(p))$ is a mixed strategy *Oligopoly Equilibrium* (OE) if the function $x^*(p) \in W(p)$ for every $p \in [0, R]^I$ and

$$\int_{[0,R]^I} \Pi_s(p_s, p_{-s}, x^*(p_s, p_{-s})) \quad d\left(\mu_s^*(p_s) \times \mu_{-s}^*(p_{-s})\right)$$

$$\geq \int_{[0,R]^I} \Pi_s(p_s, p_{-s}, x^*(p_s, p_{-s})) d\left(\mu_s(p_s) \times \mu_{-s}^*(p_{-s})\right)$$

for all s and $\mu_s \in \mathcal{B}^{I_s}$.

Therefore, a mixed strategy OE simply requires that there is no profitable deviation to a different probability measure for each oligopolist.

14

Example 3 (continued) We now show that the following strategy profile is a mixed strategy OE for the above game when $\epsilon \to 0$ (a mixed strategy OE also exists when $\epsilon > 0$, but its structure is more complicated and less informative):

$$
\mu_1(p) = \begin{cases} 0 & 0 \le p \le R(1-\delta), \\ 1 - \frac{(1-\delta)R}{p} & R(1-\delta) \le p < R, \\ 1 & \text{otherwise,} \end{cases}
$$

$$
\mu_2(p) = \begin{cases} 0 & 0 \le p \le R(1-\delta), \\ \frac{1}{\delta} - \frac{(1-\delta)R}{\delta p} & R(1-\delta) \le p \le R, \\ 1 & \text{otherwise.} \end{cases}
$$

Notice that μ_1 has an atom equal to $1 - \delta$ at R. To verify that this profile is a mixed strategy OE, let μ' be the density of μ, with the convention that $\mu' = \infty$ when there is an atom at that point. Let $M_i = \{p \mid \mu_i'(p) > 0\}$. To establish that (μ_1, μ_2) is a mixed strategy equilibrium, it suffices to show that the expected payoff to player i is constant for all $p_i \in M_i$ when the other player chooses p_{-i} according to μ_{-i} (see [24]). These expected payoffs are

$$
\bar{\Pi}\left(p_i \mid \mu_{-i}\right) \equiv \int_0^R \Pi_i(p_i, p_{-i}, x\left(p_i, p_{-i}\right)) d\mu_{-i}\left(p_{-i}\right),
$$

which are constant for all $p_i \in M_i$ for $i = 1, 2$. This establishes that (μ_1, μ_2) is a mixed strategy OE.

The next proposition shows that a mixed strategy equilibrium always exists.

Proposition 8 Let Assumption 1 hold. Then the price competition game has a mixed strategy OE, $(\mu^{OE}, x^{OE}(p))$.

The proof of this proposition is long and not directly related the rest of the argument, so it is given in Appendix B.

We next provide an explicit characterization of pure strategy OE. Though of also independent interest, these results are most useful for us to quantify the efficiency loss of oligopoly in the next section.

The following lemma shows that an equivalent to Lemma 1 (which required real-valued latency functions) also holds with more general latency functions at the pure strategy OE.

Lemma 4 Let Assumption 1 hold. If (p^{OE}, x^{OE}) is a pure strategy OE, then

$$
l_i(x_i^{OE}) + p_i^{OE} = \min_j \{l_j(x_j^{OE}) + p_j^{OE}\}, \qquad \forall\ i \text{ with } x_i^{OE} > 0, \tag{13}
$$

$$
l_i(x_i^*) + p_i^{OE} \le R, \qquad \forall\ i \text{ with } x_i^{OE} > 0, \tag{14}
$$

$$
\sum_{i=1}^I x_i^{OE} \le d, \tag{15}
$$

with $\sum_{i=1}^{I} x_i^{OE} = d$ if $\min_j \{ l_j(x_j^{OE}) + p_j \} < R$.

Proof. Let (p^{OE}, x^{OE}) be an OE. Since $x^{OE} \in W(p^{OE})$, conditions (14) and (15) follow by the definition of a WE. Consider condition (13). Assume that there exist some $i, j \in \mathcal{I}$ with $x_i^{OE} > 0$, $x_j^{OE} > 0$ such that

$$l_i(x_i^{OE}) + p_i^{OE} < l_j(x_j^{OE}) + p_j^{OE}.$$

Using the optimality conditions for a WE [cf. Eq. (2)], this implies that $x_i^{OE} = b_{C_i}$. Consider changing p_i^{OE} to $p_i^{OE} + \epsilon$ for some $\epsilon > 0$. By checking the optimality conditions, we see that we can choose ϵ sufficiently small such that $x^{OE} \in W(p_i^{OE} + \epsilon, p_{-i}^{OE})$. Hence service provider that owns link i can deviate to $p_i^{OE} + \epsilon$ and increase its profits, contradicting the fact that (p^{OE}, x^{OE}) is an OE. Finally, assume to arrive at the contradiction that $\min_j \{ l_j(x_j^{OE}) + p_j \} < R$ and $\sum_{i=1}^{I} x_i^{OE} < d$. Using the optimality conditions for a WE [Eq. (2) with $\lambda = 0$ since $\sum_{i=1}^{I} x_i^{OE} < d$], this implies that we must have $x_i^{OE} = b_{C_i}$ for some i. With a similar argument to above, a deviation to $p_i^{OE} + \epsilon$ keeps x^{OE} as a WE, and is more profitable, completing the proof. **Q.E.D.**

We need the following additional assumption for our price characterization.

Assumption 2 Given a pure strategy OE (p^{OE}, x^{OE}), if for some $i \in \mathcal{I}$ with $x_i^{OE} > 0$, we have $l_i(x_i^{OE}) = 0$, then $\mathcal{I}_s = \{i\}$.

Note that this assumption is automatically satisfied if all latency functions are strictly increasing or if all service providers own only one link.

Lemma 5 Let (p^{OE}, x^{OE}) be a pure strategy OE. Let Assumptions 1 and 2 hold. Let Π_s denote the profit of service provider s at (p^{OE}, x^{OE}).

(a) If $\Pi_{s'} > 0$ for some $s' \in \mathcal{S}$, then $\Pi_s > 0$ for all $s \in \mathcal{S}$.

(b) If $\Pi_s > 0$ for some $s \in \mathcal{S}$, then $p_j^{OE} x_j^{OE} > 0$ for all $j \in \mathcal{I}_s$.

Proof.

(a) For some $j \in \mathcal{I}_{s'}$, define $K = p_j^{OE} + l_j(x_j^{OE})$, which is positive since $\Pi_{s'} > 0$. Assume $\Pi_s = 0$ for some s. For $k \in \mathcal{I}_s$, consider the price $\bar{p}_k = K - \epsilon > 0$ for some small $\epsilon > 0$. It can be seen that at the price vector $(\bar{p}_k, p_{-k}^{OE})$, the corresponding WE link flow would satisfy $\bar{x}_k > 0$. Hence, service provider s has an incentive to deviate to $\bar{p}_k$ at which he will make positive profit, contradicting the fact that (p^{OE}, x^{OE}) is a pure strategy OE.

(b) Since $\Pi_s > 0$, we have $p_m^{OE} x_m^{OE} > 0$ for some $m \in \mathcal{I}_s$. By Assumption 2, we can assume without loss of generality that $l_m(x_m^{OE}) > 0$ (otherwise, we are done). Let $j \in \mathcal{I}_s$ and assume to arrive at a contradiction that $p_j^{OE} x_j^{OE} = 0$. The profit of service provider s at the pure strategy OE can be written as

$$\Pi_s = \bar{\Pi}_s + p_m^{OE} x_m^{OE},$$

where $\bar{\Pi}_s$ denotes the profits from links other than m and j. Let $p_m^{OE} = K - l_m(x_m^{OE})$ for some K. Consider changing the prices p_m^{OE} and p_j^{OE} such that the new profit is

$$\tilde{\Pi}_s = \bar{\Pi}_s + (K - l_m(x_m^{OE} - \epsilon))(x_m^{OE} - \epsilon) + \epsilon(K - l_j(\epsilon)).$$

Note that ϵ units of flow are moved from link m to link j such that the flows of other links remain the same at the new WE. Hence, the change in the profit is

$$\tilde{\Pi}_s - \Pi_s = (l_m(x_m^{OE}) - l_m(x_m^{OE} - \epsilon))x_m^{OE} + \epsilon(l_m(x_m^{OE} - \epsilon) - l_j(\epsilon))).$$

Since $l_m(x_m^{OE}) > 0$, ϵ can be chosen sufficiently small such that the above is strictly positive, contradicting the fact that (p^{OE}, x^{OE}) is an OE. **Q.E.D.**

The following example shows that Assumption 2 cannot be dispensed with for part (b) of this lemma.

Example 4 Consider a three link network with two providers, where provider 1 owns links 1 and 3 and provider 2 owns link 2. Let the total flow be $d = 1$ and the reservation utility be $R = 1$. Assume that the latency functions are given by

$$l_1(x_1) = 0, \quad l_2(x_2) = x_2, \quad l_3(x_3) = ax_3,$$

for some $a > 0$. Any price vector $(p_1, p_2, p_3) = (2/3, 1/3, b)$ with $b \geq 2/3$ and $(x_1, x_2, x_3) = (2/3, 1/3, 0)$ is a pure strategy OE, so $p_3 x_3 = 0$ contrary to part (b) of the lemma. To see why this is an equilibrium, note that provider 2 is clearly playing a best response. Moreover, in this allocation $\Pi_1 = 4/9$. We can represent any deviation of provider 1 by

$$(p_1, p_3) = (2/3 - \delta, 2/3 - a\epsilon - \delta),$$

for two scalars ϵ and δ, which will induce a WE of $(x_1, x_2, x_3) = (2/3 + \delta - \epsilon, 1/3 - \delta, \epsilon)$. The corresponding profit of provider 1 at this deviation is $\Pi_1 = 4/9 - \delta^2 < 4/9$, establishing that provider 1 is also playing a best response and we have a pure strategy OE.

We next establish that under an additional assumption, a pure strategy OE will never be at a point of non-differentiability of the latency functions.

Assumption 3 There exists some $s \in \mathcal{S}$ such that l_i is real-valued and continuously differentiable for all $i \in \mathcal{I}_s$.

Lemma 6 Let (p^{OE}, x^{OE}) be an OE with $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$ and $p_i^{OE} x_i^{OE} > 0$ for some i. Let Assumptions 1, 2 and 3 hold. Then

$$l_i^+(x_i^{OE}) = l_i^-(x_i^{OE}), \qquad \forall\, i \in \mathcal{I},$$

where $l_i^+(x_i^{OE})$ and $l_i^-(x_i^{OE})$ are the right and left derivatives of the function l_i at x_i^{OE} respectively.

Since the proof of this lemma is long, it is given in Appendix C. Note that Assumption 3 cannot be dispensed with in this lemma. This is illustrated in the next example.

Example 5 Consider a two link network. Let the total flow be $d = 1$ and the reservation utility be $R = 2$. Assume that the latency functions are given by

$$l_1(x) = l_2(x) = \begin{cases} 0 & \text{if } 0 \le x \le \frac{1}{2} \\ 2\left(x - \frac{1}{2}\right) & \text{otherwise.} \end{cases}$$

It can be verified that the vector $(p_1^{OE}, p_2^{OE}) = (1, 1)$, with $(x_1^{OE}, x_2^{OE}) = (1/2, 1/2)$ is a pure strategy OE, and is at a point of non-differentiability for both latency functions.

We next provide an explicit characterization of the OE prices, which is essential in our efficiency analysis in Section 5. The proof is given in Appendix D.

Proposition 9 Let (p^{OE}, x^{OE}) be an OE such that $p_i^{OE} x_i^{OE} > 0$ for some $i \in \mathcal{I}$. Let Assumptions 1, 2, and 3 hold.

a) Assume that $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$. Then, for all $s \in \mathcal{S}$ and $i \in \mathcal{I}_s$, we have

$$p_i^{OE} = \begin{cases} x_i^{OE} l_i'(x_i^{OE}), & \text{if } l_j'(x_j^{OE}) = 0 \text{ for some } j \notin \mathcal{I}_s, \\ x_i^{OE} l_i'(x_i^{OE}) + \frac{\sum_{j \in \mathcal{I}_s} x_j^{OE}}{\sum_{j \notin \mathcal{I}_s} \frac{1}{l_j'(x_j^{OE})}}, & \text{otherwise.} \end{cases} \tag{16}$$

b) Assume that $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} = R$. Then, for all $s \in \mathcal{S}$ and $i \in \mathcal{I}_s$, we have

$$p_i^{OE} \ge x_i^{OE} l_i^-(x_i^{OE}). \tag{17}$$

Moreover, if there exists some $i \in \mathcal{I}$ such that $\mathcal{I}_s = \{i\}$ for some $s \in \mathcal{S}$, then

$$p_i^{OE} \le x_i^{OE} l_i^+(x_i^{OE}) + \frac{x_i^{OE}}{\sum_{j \ne 1} \frac{1}{l_j^-(x_j^{OE})}}. \tag{18}$$

If the latency functions l_i are all real-valued and continuously differentiable, then analysis of Karush-Kuhn-Tucker conditions for oligopoly problem [Eq. (78) in Appendix D] immediately yields the following result:

Corollary 1 Let (p^{OE}, x^{OE}) be an OE such that $p_i^{OE} x_i^{OE} > 0$ for some $i \in \mathcal{I}$. Let Assumptions 1 and 2 hold. Assume also that l_i is real-valued and continuously differentiable for all i. Then, for all $s \in \mathcal{S}$ and $i \in \mathcal{I}_s$, we have

$$p_i^{OE} = \begin{cases} x_i^{OE} l_i'(x_i^{OE}), & \text{if } l_j'(x_j^{OE}) = 0 \text{ for some } j \notin \mathcal{I}_s, \\ \min \left\{ R - l_i(x_i^{OE}) \ , \ x_i^{OE} l_i'(x_i^{OE}) + \frac{\sum_{j \in \mathcal{I}_s} x_j^{OE}}{\sum_{j \notin \mathcal{I}_s} \frac{1}{l_j'(x_j^{OE})}} \right\}, & \text{otherwise.} \end{cases} \tag{19}$$

This corollary also implies that in the two link case with real-valued and continuously differentiable latency functions and with minimum effective cost less than R, the OE prices are

$$p_i^{OE} = x_i^{OE} (l_1'(x_1^{OE}) + l_2'(x_2^{OE})) \tag{20}$$

as claimed in the Introduction.

18

5 Efficiency of Oligopoly Equilibria

In this section, we study the efficiency properties of oligopoly pricing. We take as our measure of efficiency the ratio of the social surplus of the equilibrium flow allocation to the social surplus of the social optimum, $\mathbb{S}(x^*)/\mathbb{S}(x^S)$, where x^* refers to the monopoly or the oligopoly equilibrium [cf. Eq. (8)]. In Section 3, we showed that the flow allocation at a monopoly equilibrium is a social optimum. Hence, in congestion games with monopoly pricing, there is no efficiency loss. The following example shows that this is not necessarily the case in oligopoly pricing.

Example 6 Consider a two link network. Let the total flow be $d = 1$ and the reservation utility be $R = 1$. The latency functions are given by

$$l_1(x) = 0, \qquad l_2(x) = \frac{3}{2}x.$$

The unique social optimum for this example is $x^S = (1, 0)$. The unique ME (p^{ME}, xME) is $x^{ME} = (1, 0)$ and $p^{ME} = (1, 1)$. As expected, the flow allocations at the social optimum and the ME are the same. Next consider a duopoly where each of these links is owned by a different provider. Using Corollary 1 and Lemma 4, it follows that the flow allocation at the OE, x^{OE}, satisfies

$$l_1(x_1^{OE}) + x_1^{OE}[l_1'(x_1^{OE}) + l_2'(x_2^{OE})] = l_2(x_2^{OE}) + x_2^{OE}[l_1'(x_1^{OE}) + l_2'(x_2^{OE})].$$

Solving this together with $x_1^{OE} + x_2^{OE} = 1$ shows that the flow allocation at the unique oligopoly equilibrium is $x^{OE} = (2/3, 1/3)$. The social surplus at the social optimum, the monopoly equilibrium, and the oligopoly equilibrium are given by 1, 1, and 5/6, respectively.

Before providing a more thorough analysis of the efficiency properties of the OE, the next proposition proves that, as claimed in the Introduction and suggested by Example 6, a change in the market structure from monopoly to duopoly in a two link network typically reduces efficiency.

Proposition 10 Consider a two link network where each link is owned by a different provider. Let Assumption 1 hold. Let (p^{OE}, x^{OE}) be a pure strategy OE such that $p_i^{OE} x_i^{OE} > 0$ for some $i \in \mathcal{I}$ and $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$. If $l_1'(x_1^{OE})/x_1^{OE} \neq l_2'(x_2^{OE})/x_2^{OE}$, then $\mathbb{S}(x^{OE})/\mathbb{S}(x^S) < 1$.

Proof. Combining the OE prices with the WE conditions, we have

$$l_1(x_1^{OE}) + x_1^{OE}(l_1'(x_1^{OE}) + l_2'(x_2^{OE})) = l_2(x_2^{OE}) + x_2^{OE}(l_1'(x_1^{OE}) + l_2'(x_2^{OE})),$$

where we use the fact that $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$. Moreover, we can use optimality conditions (7) to prove that a vector $(x_1^S, x_2^S) > 0$ is a social optimum if and only if

$$l_1(x_1^S) + x_1^S l_1'(x_1^S) = l_2(x_2^S) + x_2^S l_2'(x_2^S),$$

(see also the proof of Lemma 4). Since $l_1'(x_1^{OE})/x_1^{OE} \neq l_2'(x_2^{OE})/x_2^{OE}$, the result follows. **Q.E.D.**

We next quantify the efficiency of oligopoly equilibria by providing a tight bound on the efficiency loss in congestion games with oligopoly pricing. As we have shown in Section 4, such games do not always have a pure strategy OE. In the following, we first provide bounds on congestion games that have pure strategy equilibria. We next study efficiency properties of mixed strategy equilibria.

5.1 Pure Strategy Equilibria

We consider price competition games that have pure strategy equilibria (this set includes, but is larger than, games with linear latency functions, see Section 4). We consider latency functions that satisfy Assumptions 1, 2, and 3. Let $\mathcal{L}_I$ denote the set of latency functions $\{l_i\}_{i \in \mathcal{I}}$ such that the associated congestion game has a pure strategy OE and the individual l_i's satisfy Assumptions 1, 2, and 3.[7] Given a parallel link network with I links and latency functions $\{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I$, let $\overrightarrow{OE}(\{l_i\})$ denote the set of flow allocations at an OE. We define the efficiency metric at some $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ as

$$r_I(\{l_i\}, x^{OE}) = \frac{R \sum_{i=1}^I x_i^{OE} - \sum_{i=1}^I l_i(x_i^{OE}) x_i^{OE}}{R \sum_{i=1}^I x_i^S - \sum_{i=1}^I l_i(x_i^S) x_i^S}, \tag{21}$$

where x^S is a social optimum given the latency functions $\{l_i\}_{i \in \mathcal{I}}$ and R is the reservation utility. In other words, our efficiency metric is the ratio of the social surplus in an equilibrium relative to the surplus in the social optimum. Following the literature on the "price of anarchy", in particular [19], we are interested in the worst performance in an oligopoly equilibrium, so we look for a lower bound on

$$\inf_{\{l_i\} \in \mathcal{L}_I} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_I(\{l_i\}, x^{OE}).$$

We first prove two lemmas, which reduce the set of latency functions that need to be considered in bounding the efficiency metric. The next lemma allows us to use the oligopoly price characterization given in Proposition 9

Lemma 7 Let (p^{OE}, x^{OE}) be a pure strategy OE such that $p_i^{OE} x_i^{OE} = 0$ for all $i \in \mathcal{I}$. Then x^{OE} is a social optimum.

Proof. We first show that $l_i(x_i^{OE}) = 0$ for all $i \in \mathcal{I}$. Assume that $l_j(x_j^{OE}) > 0$ for some $j \in \mathcal{I}$. This implies that $x_j^{OE} > 0$ and therefore $p_j^{OE} = 0$. Since $l_j(x_j^{OE}) > 0$, it follows by Lemma 2 that for all $x \in W(p)$, we have $x_j = x_j^{OE}$. Consider increasing p_j^{OE} to some small $\epsilon > 0$. By the upper semicontinuity of $W(p)$, it follows that for all $x \in W(\epsilon, p_{-j}^{OE})$, we have $|x_j - x_j^{OE}| < \delta$ for some $\delta > 0$. Moreover, by Lemma 2, we have, for all $x \in W(\epsilon, p_{-j}^{OE})$, $x_i \geq x_i^{OE}$ for all $i \neq j$. Hence, the profit of the provider that owns

---

[7]More explicitly, Assumption 2 implies that if any OE (p^{OE}, x^{OE}) associated with $\{l_i\}_{i \in \mathcal{I}}$ has $x_i^{OE} > 0$ and $l_i(x_i^{OE}) = 0$, then $\mathcal{I}_s = \{i\}$.

link j is strictly higher at price vector (ϵ, p^{OE}_{-j}) than at p^{OE}, contradicting the fact that (p^{OE}, x^{OE}) is an OE.

Clearly $x^{OE}_j > 0$ for some j and hence $\min_{i \in \mathcal{I}} \{p^{OE}_i + l_i(x^{OE}_i)\} = p^{OE}_j + l_j(x^{OE}_j) = 0$, which implies by Lemma 4 that $\sum_{i \in \mathcal{I}} x^{OE}_i = d$. Using $l_i(x^{OE}_i) = 0$, and $0 \in \partial l_i(x^{OE}_i)$ for all i, we have

$$R - l_i(x^{OE}_i) - x^{OE}_i g_{l_i} = R, \qquad \forall \, i \in \mathcal{I},$$

for some $g_{l_i} \in \partial l_i(x^{OE}_i)$. Hence, x^{OE} satisfies the sufficient optimality conditions for a social optimum [cf. Eq. (7) with $\lambda^S = R$], and the result follows. **Q.E.D.**

The next lemma allows us to assume without loss of generality that $R \sum_{i=1}^{I} x^S_i - \sum_{i=1}^{I} l_i(x^S_i) x^S_i > 0$ and $\sum_{i=1}^{I} x^{OE}_i = d$ in the subsequent analysis.

Lemma 8 Let $\{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I$. Assume that

either (i) $\sum_{i=1}^{I} l_i(x^S_i) x^S_i = R \sum_{i=1}^{I} x^S_i$ for some social optimum x_s,

or (ii) $\sum_{i=1}^{I} x^{OE}_i < d$ for some $x^{OE} \in \overrightarrow{OE}(\{l_i\})$.

Then every $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ is a social optimum, implying that $r_I(\{l_i\}, x^{OE}) = 1$.

Proof. Assume that $\sum_{i=1}^{I} l_i(x^S_i) x^S_i = R \sum_{i=1}^{I} x^S_i$. Since x^S is a social optimum and every $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ is a feasible solution to the social problem [problem (6)], we have

$$0 = \sum_{i=1}^{I} (R - l_i(x^S_i)) x^S_i \geq \sum_{i=1}^{I} (R - l_i(x^{OE}_i)) x^{OE}_i, \qquad \forall \, x^{OE} \in \overrightarrow{OE}(\{l_i\}).$$

By the definition of a WE, we have $x^{OE}_i \geq 0$ and $R - l_i(x^{OE}_i) \geq p^{OE}_i \geq 0$ (where p^{OE}_i is the price of link i at the OE) for all i. This combined with the preceding relation shows that x^{OE} is a social optimum.

Assume next that $\sum_{i=1}^{I} x^{OE}_i < d$ for some $x^{OE} \in \overrightarrow{OE}(\{l_i\})$. Let p^{OE} be the associated OE price. Assume that $p^{OE}_j x^{OE}_j > 0$ for some $j \in \mathcal{I}$ (otherwise we are done by Lemma 7). Since $\sum_{i=1}^{I} x^{OE}_i < d$, we have by Lemma 4 that $\min_{j \in \mathcal{I}} \{p_j + l_j(x^{OE}_j)\} = R$. Moreover, by Lemma 5, it follows that $p_i x^{OE}_i > 0$ for all $i \in \mathcal{I}$. Hence, for all $s \in \mathcal{S}$, $((p^{OE}_i)_{i \in \mathcal{I}_s}, x^{OE})$ is an optimal solution of the problem

$$\begin{aligned}
\text{maximize}_{((p_i)_{i \in \mathcal{I}_s}, x)} \quad & \sum_{i \in \mathcal{I}_s} p_i x_i \\
\text{subject to} \quad & p_i + l_i(x_i) = R, \qquad \forall \, i \in \mathcal{I}_s, \\
& p^{OE}_i + l_i(x_i) = R, \qquad \forall \, i \notin \mathcal{I}_s, \\
& \sum_{i=1}^{I} x^{OE}_i \leq d.
\end{aligned}$$

Substituting for $(p_i)_{i \in \mathcal{I}_s}$ in the above, we obtain

$$\text{maximize}_{x \geq 0} \quad \sum_{i \in \mathcal{I}_s} (R - l_i(x_i)) x_i$$

21

$$\text{subject to} \quad x_i \in T_i, \qquad \forall \, i \notin \mathcal{I}_s,$$

$$\sum_{i=1}^{I} x_i^{OE} \leq d,$$

where $T_i = \{x_i \mid p_i^{OE} + l_i(x_i) = R\}$ is either a singleton or a closed interval. Since this is a convex problem, using the optimality conditions, we obtain

$$R - l_i(x_i^{OE}) - x_i^{OE} g_{l_i} = 0, \qquad \forall \, i \in \mathcal{I}_s, \ \forall \, s \in \mathcal{S},$$

where $g_{l_i} \in \partial l_i(x_i^{OE})$. By Eq. (7), it follows that x^{OE} is a social optimum. **Q.E.D.**

Hence, in finding a lower bound on the efficiency metric, we can restrict ourselves, without loss of generality, to latency functions $\{l_i\} \in \mathcal{L}_I$ such that $\sum_{i=1}^{I} l_i(x_i^S) x_i^S < R \sum_{i=1}^{I} x_i^S$ for some social optimum x^S, and $\sum_{i=1}^{I} x_i^{OE} = d$ for all $x^{OE} \in \overrightarrow{OE}(\{l_i\})$. By the following lemma, we can also assume that $\sum_{i=1}^{I} x_i^S = d$.

Lemma 9 For a set of latency functions $\{l_i\}_{i \in \mathcal{I}}$, let Assumption 1 hold. Let (p^{OE}, x^{OE}) be an OE and x^S be a social optimum. Then

$$\sum_{i \in \mathcal{I}} x_i^{OE} \leq \sum_{i \in \mathcal{I}} x_i^S.$$

Proof. Assume to arrive at a contradiction that $\sum_{i \in \mathcal{I}} x_i^{OE} > \sum_{i \in \mathcal{I}} x_i^S$. This implies that $x_j^{OE} > x_j^S$ for some j. We also have $l_j(x_j^{OE}) > l_j(x_j^S)$. (Otherwise, we would have $l_j(x_j^S) = l_j'(x_j^S) = 0$, which yields a contradiction by the optimality conditions (7) and the fact that $\sum_{i \in \mathcal{I}} x_i^S < d$). Using the optimality conditions (2) and (7), we obtain

$$R - l_j(x_j^{OE}) - p_j^{OE} \geq R - l_j(x_j^S) - x_j^S g_{l_j},$$

for some $g_{l_j} \in \partial l_j(x_j^S)$. Combining the preceding with $l_j(x_j^{OE}) > l_j(x_j^S)$ and $p_j^{OE} \geq x_j^{OE} l_j^-(x_j^{OE})$ (cf. Proposition 9), we see that

$$x_j^{OE} l_j^-(x_j^{OE}) < x_j^S g_{l_j},$$

contradicting $x_j^{OE} > x_j^S$ and completing the proof. **Q.E.D.**

5.1.1 Two Links

We first consider a parallel link network with two links owned by two service providers. The next theorem provides a tight lower bound on $r_2(\{l_i\}, x^{OE})$ [cf. Eq. (21)]. In the following, we assume without loss of generality that $d = 1$. Also recall that latency functions in $\mathcal{L}_2$ satisfy Assumptions 1, 2, and 3.

Theorem 1 Consider a two link network where each link is owned by a different provider. Then

$$r_2(\{l_i\}, x^{OE}) \geq \frac{5}{6}, \qquad \forall \, \{l_i\}_{i=1,2} \in \mathcal{L}_2, \ x^{OE} \in \overrightarrow{OE}(\{l_i\}), \tag{22}$$

and the bound is tight, i.e., there exists $\{l_i\}_{i=1,2} \in \mathcal{L}_2$ and $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ that attains the lower bound in Eq. (22).

Proof. The proof follows a number of steps:

Step 1: We are interested in finding a lower bound for the problem

$$\inf_{\{l_i\} \in \mathcal{L}_2} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_2(\{l_i\}, x^{OE}). \tag{23}$$

Given $\{l_i\} \in \mathcal{L}_2$, let $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ and let x^S be a social optimum. By Lemmas 8 and 9, we can assume that $\sum_{i=1}^2 x_i^{OE} = \sum_{i=1}^2 x_i^S = 1$. This implies that there exists some i such that $x_i^{OE} < x_i^S$. Since the problem is symmetric, we can restrict ourselves to $\{l_i\} \in \mathcal{L}_2$ such that $x_1^{OE} < x_1^S$. We claim

$$\inf_{\{l_i\} \in \mathcal{L}_2} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_2(\{l_i\}, x^{OE}) \geq r_{2,t}^{OE}, \tag{24}$$

where

$$r_{2,t}^{OE} = \operatorname*{minimize}_{\substack{l_i^S, \, (l_i^S)' \geq 0 \\ l_i, \, l_i' \geq 0 \\ y_i^S, \, y_i^{OE} \geq 0}} \frac{R - l_1 y_1^{OE} - l_2 y_2^{OE}}{R - l_1^S y_1^S - l_2^S y_2^S} \tag{E}$$

$$\text{subject to} \quad l_i^S \leq y_i^S (l_i^S)', \qquad\qquad i = 1, 2, \tag{25}$$

$$l_2^S + y_2^S (l_2^S)' = l_1^S + y_1^S (l_1^S)', \tag{26}$$

$$l_1^S + y_1^S (l_1^S)' \leq R, \tag{27}$$

$$\sum_{i=1}^2 y_i^S = 1, \tag{28}$$

$$l_1 \leq l_1^S, \tag{29}$$

$$l_1' = 0, \qquad \text{if } l_1^S = 0. \tag{30}$$

$$l_i \leq y_i^{OE} l_i', \qquad i = 1, 2, \tag{31}$$

$$\sum_{i=1}^2 y_i^{OE} = 1, \tag{32}$$

$$+ \{\text{Oligopoly Equilibrium Constraints}\}_t, \qquad t = 1, 2.$$

Problem (E) can be viewed as a finite dimensional problem that captures the equilibrium and the social optimum characteristics of the infinite dimensional problem given in Eq. (23). This implies that instead of optimizing over the entire function l_i, we optimize over the possible values of $l_i(\cdot)$ and $\partial l_i(\cdot)$ at the equilibrium and the social optimum, which we denote by $l_i, l_i', l_i^S, (l_i^S)'$ [i.e., $(l_i^S)'$ is a variable that represents all possible values of $g_{l_i} \in \partial l_i(y_i^S)$]. The constraints of the problem guarantee that these values satisfy the necessary optimality conditions for a social optimum and an OE. In particular, conditions (25) and (31) capture the convexity assumption on $l_i(\cdot)$ by relating the values l_i, l_i' and $l_i^S, (l_i^S)'$ [note that the assumption $l_i(0) = 0$ is essential here]. Condition (26) is the

23

optimality condition for the social optimum. Condition (29) captures the nondecreasing assumption on the latency functions; since we are considering $\{l_i\}$ such that $x_1^{OE} < x_1^S$, we must have $l_1 \leq l_1^S$. Condition (30) captures the relation of $l_1(\cdot)$ and $l_1'(\cdot)$ at x_1^{OE}; since $x_1^{OE} < x_1^S$, the fact that $l_1(x_1^{OE}) = l_1(x_1^S) = 0$ implies, by Assumption 1, that 0 is the unique element of $\partial l_1(x_1^{OE})$. Finally, the last set of constraints are the necessary conditions for a pure strategy OE. These are written separately for $t = 1, 2$, for the two cases characterized in Proposition 9, giving us two bounds, which we will show to be equal.

More explicitly, the Oligopoly Equilibrium constraints are given by:

For $t = 1$: [corresponding to a lower bound for pure strategy OE, (p^{OE}, y^{OE}), with $\min_j\{p_j^{OE} + l_j(y_j^{OE})\} < R$],

$$
\begin{aligned}
l_1 + y_1^{OE}[l_1' + l_2'] &= l_2 + y_2^{OE}[l_1' + l_2'], \\
l_1 + y_1^{OE}[l_1' + l_2'] &\leq R,
\end{aligned}
\tag{33}
$$

where $l_1' = l_1'(y_1^{OE})$, $l_2' = l_2'(y_2^{OE})$ [cf. Eq. (16)].

For $t = 2$: [corresponding to a lower bound for pure strategy OE, (p^{OE}, y^{OE}), with $\min_j\{p_j^{OE} + l_j(y_j^{OE})\} = R$],

$$
\begin{aligned}
R - l_2 &\geq y_2^{OE} l_2', \\
R - l_1 &\leq y_1^{OE}[l_1' + l_2'],
\end{aligned}
\tag{34}
$$

where $l_1' = l_1^+(y_1^{OE})$ and $l_2' = l_2^-(y_2^{OE})$ [cf. Eqs. (17), (18)]. We will show in Step 4 that $r_{2,1}^{OE} = r_{2,2}^{OE}$.

Note that given any feasible solution of problem (23), we have a feasible solution for problem (E) with the same objective function value. Therefore, the optimum value of problem (E) is indeed a lower bound on the optimum value of problem (23).

Step 2: Let $(l_i^S, y_i^S)_{i=1,2}$ satisfy Eqs. (25)-(28). We show that

$$
l_1^S y_1^S + l_2^S y_2^S < R.
\tag{35}
$$

Using Eqs. (26), (27), and (28), we obtain

$$
l_1^S y_1^S + l_2^S y_2^S + (y_1^S)^2 (l_1^S)' + (y_2^S)^2 (l_2^S)' \leq R.
$$

If $(y_1^S)^2 (l_1^S)' + (y_2^S)^2 (l_2^S)' > 0$, then the result follows. If $(y_1^S)^2 (l_1^S)' + (y_2^S)^2 (l_2^S)' = 0$, then we have using Eq. (25) that $l_i^S = 0$ for all i, again showing the result.

Next, let $(l_i, y_i^{OE})_{i=1,2}$ satisfy Eq. (32) and one of the Oligopoly Equilibrium constraints [i.e., Eqs. (33) or (34)]. Using a similar argument, we can show that

$$
l_1 y_1^{OE} + l_2 y_2^{OE} < R.
\tag{36}
$$

Step 3: To solve problem (E), we first relax the last constraint [Eq. (30)]. Let $(\bar{l}_i^S, (\bar{l}_i^S)', \bar{l}_i, \bar{l}_i', \bar{y}_i^S, \bar{y}_i^{OE})$ denote the optimal solution of the relaxed problem [problem (E) without constraint (30)]. We show that $\bar{l}_i^S = 0$ for $i = 1, 2$.

Assign the Lagrange multipliers $\mu_i^S, \lambda^S, \gamma^S$ to Eqs. (25), (26), (27), respectively, and multiplier θ^S to Eq. (28). Using the first order optimality conditions, we obtain

$$\bar{y}_2^S \frac{(R - \bar{l}_1 \bar{y}_1^{OE} - \bar{l}_2 \bar{y}_2^{OE})}{(R - \bar{l}_1^S \bar{y}_1^S - \bar{l}_2^S \bar{y}_2^S)^2} + \mu_2^S + \lambda^S \quad = 0 \quad \text{if } \bar{l}_2^S > 0 \tag{37}$$
$$\geq 0 \quad \text{if } \bar{l}_2^S = 0,$$

$$-\mu_2^S \bar{y}_2^S + \lambda^S \bar{y}_2^S \quad = 0 \quad \text{if } (\bar{l}_2^S)' > 0 \tag{38}$$
$$\geq 0 \quad \text{if } (\bar{l}_2^S)' = 0,$$

$$-\mu_1^S \bar{y}_1^S - \lambda^S \bar{y}_1^S + \gamma^S \bar{y}_1^S \quad = 0 \quad \text{if } (\bar{l}_1^S)' > 0 \tag{39}$$
$$\geq 0 \quad \text{if } (\bar{l}_1^S)' = 0.$$

$$\bar{l}_1^S \frac{(R - \bar{l}_1 \bar{y}_1^{OE} - \bar{l}_2 \bar{y}_2^{OE})}{(R - \bar{l}_1^S \bar{y}_1^S - \bar{l}_2^S \bar{y}_2^S)^2} - \mu_1^S (\bar{l}_1^S)' - \lambda^S (\bar{l}_1^S)' + \gamma^S (\bar{l}_1^S)' + \theta^S \quad = 0 \quad \text{if } \bar{y}_1^S > 0 \tag{40}$$
$$\geq 0 \quad \text{if } \bar{y}_1^S = 0,$$

$$\bar{l}_2^S \frac{(R - \bar{l}_1 \bar{y}_1^{OE} - \bar{l}_2 \bar{y}_2^{OE})}{(R - \bar{l}_1^S \bar{y}_1^S - \bar{l}_2^S \bar{y}_2^S)^2} - \mu_2^S (\bar{l}_2^S)' + \lambda^S (\bar{l}_1^S)' + \theta^S \quad = 0 \quad \text{if } \bar{y}_2^S > 0 \tag{41}$$
$$\geq 0 \quad \text{if } \bar{y}_2^S = 0,$$

We first show that $\bar{l}_2^S = 0$. If $\bar{y}_2^S = 0$ or $(\bar{l}_2^S)' = 0$, we are done by Eq. (25). Assume that $\bar{y}_2^S > 0$ and $(\bar{l}_2^S)' > 0$. By Eq. (38), this implies that $\lambda^S = \mu_2^S \geq 0$. We claim that in this case Eq. (37) cannot be equal to 0. Assume to arrive at a contradiction that it is. Using Step 2 and the fact that $\bar{y}_2^S > 0$, we have $\mu_2^S + \lambda^S < 0$, which is a contradiction and shows that Eq. (37) is strictly positive. This establishes that $\bar{l}_2^S = 0$.

We next show that $\bar{l}_1^S = 0$. If $\bar{y}_1^S = 0$ or $(\bar{l}_1^S)' = 0$, we are done by Eq. (25). Assume that $\bar{y}_1^S > 0$ and $(\bar{l}_1^S)' > 0$. Assume to arrive at a contradiction that $\bar{l}_1^S > 0$. Since $\bar{l}_2^S = 0$, by Eq. (26), we have that $\bar{y}_1^S (\bar{l}_1^S)' < \bar{y}_2^S (\bar{l}_2^S)'$, and therefore both $\bar{y}_2^S > 0$ and $(\bar{l}_2^S)' > 0$. Since $\bar{l}_2^S = 0$, these imply that Eq. (25) is slack, and hence $\mu_2^S = 0$. By Eq. (38), this shows $\lambda^S \bar{y}_2^S = 0$, and hence $\lambda^S = 0$. Furthermore, since $\mu_2^S = \lambda^S = 0$, from Eq. (41) we have $\theta^S = 0$, and from Eq. (39), we have $\mu_1^S = \gamma^S$. Plugging this all in Eq. (40), we obtain a contradiction, showing that $\bar{l}_1^S = 0$.

Step 4: Since $\bar{l}_1^S = 0$, in view of Eq. (29), we have $\bar{l}_1 = 0$. Hence we can impose the extra constraint $l_1 = 0$ in problem (E) without changing the optimal function value, and we can rewrite the constraint in Eq. (30) as $l_1' = 0$. Using in addition $\bar{l}_2^S = 0$, we see that for $t = 1, 2$,

$$r_{2,t}^{OE} \geq \min_{\substack{l_2, l_2' \\ y_2^{OE}, \, y_2^{OE} \geq 0}} \quad 1 - \frac{l_2 y_2^{OE}}{R}$$
$$\text{subject to} \quad l_2 \leq y_2^{OE} l_2',$$

25

$$l_2 + y_2^{OE} l_2' \leq R,$$
$$y_1^{OE} l_2' \geq R.$$
$$\sum_{i=1}^{2} y_i^{OE} = 1,$$

which follows because any vector (y_i^{OE}, l_i, l_i') with $l_1 = 0$ and $l_1' = 0$ that satisfies Eqs. (33) or (34) is a feasible solution to the above problem. It is straightforward to show that the optimal solution of this problem is $(\bar{l}_2, \bar{l}_2', \bar{y}_1^{OE}, \bar{y}_2^{OE}) = (\frac{R}{2}, \frac{3R}{2}, \frac{2}{3}, \frac{1}{3})$, and therefore it follows that $r_{2,1}^{OE} = r_{2,2}^{OE} = 5/6$. By Eq. (24), this implies that

$$\inf_{\{l_i\} \in \mathcal{L}_2} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_2(\{l_i\}, x^{OE}) \geq \frac{5}{6}.$$

We next show that this bound is tight. Consider the latency functions $l_1(x) = 0$, and $l_2(x) = \frac{3}{2}x$. As shown in Example 6, the corresponding OE is $x^{OE} = (\frac{2}{3}, \frac{1}{3})$, and the social optimum is $x^S = (1, 0)$. Hence, the efficiency metric for these latency functions is $r_2(\{l_i\}, x^{OE}) = 5/6$, thus showing that

$$\min_{\{l_i\}_{i=1,2} \in \mathcal{L}_2} \min_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_2(\{l_i\}, x^{OE}) = \frac{5}{6}.$$

Q.E.D.

5.1.2 General Case

We next consider the general case where we have a parallel link network with I links and S service providers, where provider s owns a set of links $\mathcal{I}_s \subset \mathcal{I}$. It can be seen by augmenting a two link network with links that have latency functions

$$l(x) = \begin{cases} 0 & \text{if } x = 0, \\ \infty & \text{otherwise,} \end{cases}$$

that the lower bound in the general network case can be no higher than $\frac{5}{6}$. However, this is a degenerate example in the sense that at the OE, the flows of the links with latency functions given above are equal to 0. We next give an example of an I link network which has positive flows on all links at the OE and an efficiency metric of 5/6.

Example 7 Consider an I link network. Let the total flow be $d = 1$ and the reservation utility be $R = 1$. The latency functions are given by

$$l_1(x) = 0, \qquad l_i(x) = \frac{3}{2}(I-1)x, \quad i = 2, \ldots, I.$$

The unique social optimum for this example is $x^S = [1, 0, \ldots, 0]$. It can be seen that the flow allocation at the unique OE is

$$x^{OE} = \left[\frac{2}{3}, \frac{1}{3(I-1)}, \ldots, \frac{1}{3(I-1)} \right].$$

26

Hence, the efficiency metric for this example is

$$r_I(\{l_i\}, x^{OE}) = \frac{5}{6}.$$

The next theorem generalizes Theorem 1 to a parallel link network with $I \geq 2$ links.

Theorem 2 Consider a general parallel link network with I links and S service providers, where provider s owns a set of links $\mathcal{I}_s \subset \mathcal{I}$. Then

$$r_I(\{l_i\}, x^{OE}) \geq \frac{5}{6}, \qquad \forall \, \{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I, \; x^{OE} \in \overrightarrow{OE}(\{l_i\}), \tag{42}$$

and the bound is tight, i.e., there exists $\{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I$ and $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ that attains the lower bound in Eq. (42).

Proof. The proof again follows a number of steps:

Step 1: Consider the problem

$$\inf_{\{l_i\} \in \mathcal{L}_I} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_I(\{l_i\}, x^{OE}). \tag{43}$$

Given $\{l_i\} \in \mathcal{L}_2$, let $x^{OE} \in \overrightarrow{OE}(\{l_i\})$ and let x^S be a social optimum. By Lemmas 8 and 9, we can assume without loss of generality that $\sum_{i=1}^{2} x_i^{OE} = \sum_{i=1}^{2} x_i^S = 1$. Hence there exists some i such that $x_i^{OE} < x_i^S$. Without loss of any generality, we restrict ourselves to the set of latency functions $\{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I$ such that $x_1^{OE} < x_1^S$. Similar to the proof of Proposition 1, it can be seen that Problem (43) can be lower bounded by the optimum value of the following finite dimensional problem:

$$r_{I,t}^{OE} = \; \underset{\substack{l_i^S, (l_i^S)' \geq 0 \\ l_i, l_i' \geq 0 \\ y_i^S, y_i^{OE} \geq 0 \\ \mathcal{I}_s \subset \mathcal{I}}}{\text{minimize}} \; \frac{R - \sum_{i=1}^{n} l_i y_i^{OE}}{R - \sum_{i=1}^{n} l_i^s y_i^S} \tag{44}$$

$$\text{subject to} \quad l_i^S \leq y_i^S (l_i^S)', \qquad\qquad\qquad i = 1, \ldots, n, \tag{45}$$

$$l_i^S + y_i^S (l_i^S)' = l_1^S + y_1^S (l_1^S)', \qquad i = 2, \ldots, n, \tag{46}$$

$$l_1^S + y_1^S (l_1^S)' \leq R, \tag{47}$$

$$\sum_{i=1}^{I} y_i^S = 1 \tag{48}$$

$$l_1 \leq l_1^S, \tag{49}$$

$$l_1' = 0, \qquad\qquad\qquad \text{if } l_1^S = 0, \tag{50}$$

$$\mathcal{I}_s = \{1\} \quad \text{for some } s \qquad \text{if } l_1^S = 0, \tag{51}$$

$$+ \{\text{Oligopoly Equilibrium Constraints}\}_t, \qquad t = 1, 2.$$

27

The new feature relative to the two link case is the presence of $\mathcal{I}_s$'s as choice variables to allow a choice over possible distribution of links across service providers (with the constraint $\bigcup_s \mathcal{I}_s = \mathcal{I}$ left implicit). The oligopoly equilibrium constraints, which are again written separately for $t = 1, 2$ for the two cases in Proposition 9, depend on $\mathcal{I}_i$'s. In addition, we have added constraint (51) to impose Assumption 2 (recall that $x_1^{OE} > 0$ by Lemma 5).

Step 2: Let $(\bar{l}_i^S, (\bar{l}_i^S)', \bar{l}_i, \bar{l}_i', \bar{y}_i^S, \bar{y}_i^{OE})$ be an optimal solution of the relaxed version of the preceding problem [i.e., without constraint (50)].

Note that the constraints that involve $(l_i^S, (l_i^S)', y_i^S)$ for $i = 2, \ldots, n$ are decoupled and have the same structure as in problem (E). Therefore, by the same argument used to show $\bar{l}_2^S = 0$ in Step 3 of the proof of Proposition 1, one can show that $\bar{l}_i^S = 0$ for each $i = 2, \ldots, n$. Similarly, one can extend the same argument given in the proof of Proposition 1 to show that $\bar{l}_1^S = 0$.

Step 3: Since $\bar{l}_1^S = 0$, it follows that $\bar{l}_1 = 0$ [Eq. (49)], and we can assume without loss of generality that $\bar{l}_1' = 0$ and $\mathcal{I}_1 = \{1\}$. Therefore, using the price characterization from Proposition 9, the structure of the problem simplifies to

$$
r_{I,t}^{OE} \geq \quad \underset{\substack{l_i, l_i' \geq 0, \ i = 2, \ldots, I \\ y_i^{OE} \geq 0, \ i = 1, \ldots, I}}{\text{minimize}} \quad 1 - \frac{\sum_{i=2}^n l_i y_i^{OE}}{R}
$$

$$
\text{subject to} \quad l_i \leq y_i^{OE} l_i', \qquad i = 2, \ldots, I,
$$

$$
l_i + y_i^{OE} l_i' \leq R, \qquad i = 2, \ldots, I,
$$

$$
\frac{y_1^{OE}}{\sum_{j=2}^I \frac{1}{l_j'}} \geq R, \tag{52}
$$

$$
\sum_{i=1}^I y_i^{OE} = 1,
$$

where we have also used the fact that $\bar{l}_i^S = 0$, for $i = 2, \ldots, I$.

The first set of constraints are due to the convexity assumptions on the l_i. Similar to the two link case, the second set of constraints are due to the oligopoly equilibrium constraints (given $\bar{l}_1' = 0$ see the OE price characterization in Proposition 9).

Step 4: Let $((\bar{l}_i, \bar{l}_i')_{i=2,\ldots,I}, (\bar{y}_i^{OE})_{i=1,\ldots,I})$ denote an optimal solution of the preceding problem. Assign the Lagrange multipliers μ_i, λ_i, γ and θ consecutively to the constraints of the problem. Using the optimality conditions, we have

$$
\frac{\gamma}{\sum_{j=2}^I \frac{1}{\bar{l}_j'}} \quad = 0 \quad \text{if } \bar{y}_1^{OE} > 0 \tag{53}
$$

$$
\geq 0 \quad \text{if } \bar{y}_1^{OE} = 0,
$$

$$
-\frac{\bar{y}_i^{OE}}{R} + \mu_i + \lambda_i \quad = 0 \quad \text{if } \bar{l}_i > 0 \tag{54}
$$

$$
\geq 0 \quad \text{if } \bar{l}_i = 0,
$$

28

and for $i = 2, \ldots, I$,

$$-\mu_i \bar{y}_i^{OE} + \gamma \bar{y}_1^{OE} \frac{1/(\bar{l}_i')^2}{\left(\sum_{j=2}^{I} \frac{1}{\bar{l}_j'}\right)^2} + \lambda_i \bar{y}_i^{OE} \quad = 0 \quad \text{if } \bar{l}_i' > 0 \tag{55}$$

$$\geq 0 \quad \text{if } \bar{l}_i' = 0.$$

By feasibility, we have $\bar{y}_1^{OE} > 0$. Using symmetry, it can also be seen that $\bar{y}_i^{OE} > 0$. Hence, by Eq. (53), we have $\gamma = 0$, which together with Eq. (55) implies that $\lambda_i = \mu_i$ for all $i = 2, \ldots, I$.

Moreover, by Eq. (54), we have $\mu_i > 0$, showing that $\bar{l}_i = \bar{y}_i^{OE} \bar{l}_i'$ for each $i = 2, \ldots, I$ (i.e., the corresponding constraint is not slack). This implies that

$$\bar{l}_i + \bar{y}_i^{OE} \bar{l}_i' = R, \qquad \forall\, i = 2, \ldots, I.$$

Hence $\bar{l}_i = R/2$ and $\bar{y}_i^{OE} \bar{l}_i' = R/2$ for each $i = 2, \ldots, I$. We can also see that constraint (52) is not slack at $((\bar{l}_i, \bar{l}_i')_{i=2,\ldots,I}, (\bar{y}_i^{OE})_{i=1,\ldots,I})$ (otherwise, we can improve the objective function value), implying that

$$\frac{y_1^{OE}}{\sum_{j=2}^{I} \frac{1}{\bar{l}_j'}} = R.$$

Combining with the feasibility constraint $\bar{y}_1^{OE} = 1 - \sum_{j=2}^{I} \bar{y}_j^i$, we see that

$$1 - \frac{R}{2} \sum_{j=2}^{I} \frac{1}{\bar{l}_j'} = R \sum_{j=2}^{I} \frac{1}{\bar{l}_j'},$$

from which we obtain

$$\sum_{j=2}^{I} \bar{y}_j^{OE} = \frac{R}{2} \left(\sum_{j=2}^{I} \frac{1}{\bar{l}_j'} \right) = \frac{1}{3}.$$

Hence we have

$$r_I^{OE} = 1 - \frac{\frac{R}{2} \frac{1}{3}}{R} = \frac{5}{6},$$

showing that

$$\inf_{\{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_I(\{l_i\}, x^{OE}) \geq \frac{5}{6}.$$

Finally, Example 7 shows that the preceding bound is tight, i.e.,

$$\min_{\{l_i\}_{i \in \mathcal{I}} \in \mathcal{L}_I} \min_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_I(\{l_i\}, x^{OE}) = \frac{5}{6}.$$

Q.E.D.

Although Example 7 and Theorem 2 show that the efficiency loss in an arbitrarily large network (i.e., with $I \to \infty$) can be as high as in the two link network, the same is not the case if we start with a given I link network and replicate it n times (also increasing d to nd). In this case, it can be shown, as in analyses of the limit behavior of oligopoly models (e.g., [28], [14], [22], [39], [40]), that as $n \to \infty$, the efficiency metric tends to 1.

29

5.2   Mixed Strategy Equilibria

As we illustrated in Section 4, congestion games with latency functions that satisfy Assumptions 1, 2, and 3 may not have a pure strategy oligopoly equilibrium (cf. Example 3). Nevertheless, as shown in Proposition 8, such games always have a mixed strategy equilibrium. In this section, we discuss the efficiency properties of mixed strategy OE.

Although there has been much less interest in the efficiency properties of mixed strategy equilibria, two different types of efficiency metrics present themselves as natural candidates. The first considers the worst realization of the strategies, while the second focuses on average inefficiency across different realizations of mixed strategies. We refer to the first metric as worst-realization metric, and denote it by $\tilde{r}_I^W(\{l_i\})$, and to the second as the average metric, and denote it by $\tilde{r}_I^A(\{l_i\})$. We discuss both of these briefly.

Given a set of latency functions $\{l_i\}_{i\in\mathcal{I}}$, let $OM(\{l_i\})$ denote the set of mixed strategy equilibria. For some $\mu \in OM(\{l_i\})$, let $M_i(\mu)$ denote the support of μ_i as defined before in Example 3 [in particular, recall that $M_i = \{p \mid \mu_i'(p) > 0\}$]. Further, let

$$\overrightarrow{OE}_m(\{l_i\}, \mu) = \{x \mid x \in W(p), \text{ for some } p \text{ s.t. } p_i \in M_i(\mu) \text{ for all } i\}.$$

We define the worst-realization efficiency metric as

$$\tilde{r}_I^W(\{l_i\}) = \inf_{\{l_i\}\in\mathcal{L}_I} \inf_{\mu\in OM(\{l_i\})} \inf_{x^{OE}\in\overrightarrow{OE}_m(\{l_i\},\mu)} r_I(\{l_i\}, x^{OE}),$$

where r_I is given by Eq. (21).

Similarly, the average efficiency metric is defined as

$$\tilde{r}_I^A(\{l_i\}) = \inf_{\{l_i\}\in\mathcal{L}_I} \inf_{\mu\in OM(\{l_i\})} \int \cdots \int r_I(\{l_i\}, x^{OE}(p))d\mu_1 \cdots d\mu_S.$$

In the next example, we show that the worst-realization efficiency metric for games with no pure strategy equilibrium can be arbitrarily low.

Example 3 (continued) Consider the prices $p_1 = R$ and $p_2 = R(1-\delta)$ that satisfy $p_i \in M_i$ for the unique mixed strategy equilibrium given in Example 3 as $\epsilon \to 0$. The WE at these prices is given by

$$x^{OE} = (1 - \delta, \delta),$$

and the worst-realization efficiency metric is

$$\tilde{r}_I^W(\{l_i\}) = 1 - \delta^2,$$

which as $\delta \to 1$ goes to 0.

Next consider the average efficiency metric, $\tilde{r}_I^A(\{l_i\})$. Once again, consider the limit as $\epsilon \to 0$. Recall that we have characterized the unique mixed strategy OE above, and define $r(p_1, p_2)$ as the inefficiency at the price vector (p_1, p_2), so that

$$\tilde{r}_I^A(\{l_i\}) = \int_{(1-\delta)R}^{R} \int_{(1-\delta)R}^{R} r(p_1, p_2) d\mu_1 \times d\mu_2$$

Clearly, using the Wardrop equilibrium flow vector derived in Example 3 before, we have that

$$r\left(p_1, p_2\right) = \begin{cases} 1 & \text{if } p_1 < p_2 \\ 1 - \frac{\delta(p_1 - p_2)}{R} & \text{if } p_1 > p_2 \end{cases} ,$$

and thus,

$$\tilde{r}_I^A(\{l_i\}) \to 1 - \int_{(1-\delta)R}^R \int_{(1-\delta)R}^R \frac{\delta\left(p_1 - p_2\right)}{R} d\mu_1 \times d\mu_2$$

Thus to calculate $\tilde{r}_I^A(\{l_i\})$, we need to compute the last integral. Denoting this by A, we have

$$
\begin{aligned}
A &= \frac{\delta}{R} \int_{(1-\delta)R}^R \int_{p_2}^R \left(p_1 - p_2\right) d\mu_1 \times d\mu_2 \\
&= \frac{\delta}{R} \left[\int_{(1-\delta)R}^R \int_{(1-\delta)R}^{p_1} p_1 d\mu_2 \times d\mu_1 - \int_{(1-\delta)R}^R \int_{p_2}^R p_2 d\mu_1 \times d\mu_2 \right] \\
&= \frac{\delta}{R} \left[\int_{(1-\delta)R}^R p_1 \left(\frac{1}{\delta} - \frac{(1-\delta)R}{\delta p_1} \right) d\mu_1\left(p_1\right) - \int_{(1-\delta)R}^R (1-\delta)R d\mu_2\left(p_2\right) \right]
\end{aligned}
$$

Now recall that μ_1 has an atom equal to $1 - \delta$ at R, so

$$
\begin{aligned}
A &= \frac{\delta}{R} \left[R\left(1-\delta\right) + \int_{(1-\delta)R}^R p_1 \left(\frac{1}{\delta} - \frac{(1-\delta)R}{\delta p_1} \right) \left(\frac{(1-\delta)R}{p_1^2} \right) dp_1 - R(1-\delta) \right] \\
&= (1-\delta)^2 - (1-\delta) + (1-\delta)[\ln R - \ln((1-\delta)R)] \\
&= -(1-\delta)\delta - (1-\delta)\ln\left(1-\delta\right)
\end{aligned}
$$

It can be calculated that A reaches a maximum of approximately 0.16 for $\delta \approx 0.8$. Therefore, in this example, $\tilde{r}_I^A(\{l_i\})$ reaches $0.84 \approx 5/6$ (in fact, slightly greater than $5/6$). We conjecture, but are unable to prove, that $5/6$ is also a lower bound for the average efficiency metric, $\tilde{r}_I^A(\{l_i\})$, in mixed strategy OE. This is left as an open research question.

6   Conclusions

In this paper, we presented an analysis of competition in congested networks. We established a number of results. First, despite the potential inefficiencies of flow-routing without prices, price-setting by a monopolist always achieves the social optimum. Second, and in contrast to the monopoly result, oligopoly equilibria where multiple service providers compete are typically inefficient. Third, there is a tight bound of 5/6 on inefficiency in pure strategy oligopoly equilibria. This bound is obtained even for arbitrarily large parallel link networks. Finally, pure strategy equilibria may fail to exist when some latency functions are highly convex. Mixed strategy equilibria can lead to arbitrarily inefficient realizations.

A number of concluding comments are useful:

- Our motivating example has been the flow of information in a communication network, but our results apply equally to traffic assignment problems and oligopoly in product markets with negative externalities, congestion or snob effects (as originally suggested by Veblen [38]).

- Our analysis has been quite general, in particular, allowing for constant latencies and capacity constraints. Some of the analysis simplifies considerably when we specialize the network to increasing and real-valued (non-capacity constrained) latencies. On the other hand, the assumption that $l_i(0) = 0$ is important for our efficiency bounds. This can be relaxed to derive a slightly lower bound on the efficiency in pure strategy OE, but this significantly complicates the characterization of the form of equilibrium prices in OE (unless we assume that all latency functions are increasing rather than non-decreasing). We leave this for future work.

- One simplifying feature of our analysis is the assumption that users are "homogeneous" in the sense that the same reservation utility, R, applies to all users. It is possible to conduct a similar analysis with elastic and heterogeneous users (or traffic), but this raises a number of new and exciting challenges. For example, monopoly or oligopoly providers might want to use non-linear pricing (designed as a mechanism subject to incentive compatibility constraints of different types of users, e.g., [42]). This is an important research area for understanding equilibria in communication networks, where users often have heterogeneous quality of service requirements.

- While we have established that worst-realization efficiency metric in mixed strategy oligopoly equilibria can be arbitrarily low, a bound for average efficiency metric is an open research question.

7   Appendix A: Proof of Proposition 4

If (p^{ME}, xME) is an ME, then it is an SPE by definition. Let (p^{ME}, xME) be an SPE. Assume to arrive at a contradiction that there exists some $p \geq 0$ and $\tilde{x} \in W(p)$ such that

$$\Pi(p^{ME}, x^{ME}) < \Pi(p, \tilde{x}). \tag{56}$$

If $W(p)$ is a singleton, we immediately obtain a contradiction. Assume that $W(p)$ is not a singleton and $\sum_{i=1}^{I} x_i = \sum_{i=1}^{I} \bar{x}_i$ for all x, $\bar{x} \in W(p)$. By Lemma 2, it follows that $\Pi(p, \tilde{x}) = \Pi(p, x)$ for all $x \in W(p)$, which contradicts the fact that (p^{ME}, xME) is an SPE.

Assume finally that $W(p)$ is not a singleton and

$$\sum_{i=1}^{I} \hat{x}_i < \sum_{i=1}^{I} \bar{x}_i, \qquad \text{for some } \hat{x}, \ \bar{x} \in W(p). \tag{57}$$

For this case, we have $p_i = R$ for all $i \in \bar{\mathcal{I}}$, where

$$\bar{\mathcal{I}} = \{ i \in \mathcal{I} \mid \exists \ x, \ \bar{x} \in W(p) \text{ with } x_i \neq \bar{x}_i \},$$

[cf. Eq. (5)]. To see this, note that since $\sum_{i=1}^{I} \hat{x}_i < d$, the WE optimality conditions for $\hat{x}$ [cf. Eq. (2)] hold with $\lambda = 0$. Assume that $\tilde{p} < R$. By Lemma 2, $l_i(x_i) = 0$ for all $i \in \bar{\mathcal{I}}$. If $b_{C_i} = \infty$ for some $i \in \bar{\mathcal{I}}$, we get a contradiction by Eq. (2). Otherwise, Eq. (2) implies that $\hat{x}_i = b_{C_i}$ for all $i \in \bar{\mathcal{I}}$. Since $\hat{x}_i = \bar{x}_i$ for all $i \notin \bar{\mathcal{I}}$, this contradicts Eq. (57).

We show that given $\delta > 0$, there exists some $\epsilon > 0$ such that

$$\Pi(p^\epsilon, x^\epsilon) \geq \Pi(p, \tilde{x}) - \delta, \qquad \forall \ x^\epsilon \in W(p^\epsilon), \tag{58}$$

where

$$p_i^\epsilon = \begin{cases} p_i & i \notin \bar{\mathcal{I}}, \\ R - \epsilon & i \in \bar{\mathcal{I}}. \end{cases} \tag{59}$$

The preceding relation together with Eq. (56) contradicts the fact that (p^{ME}, xME) is an SPE, thus establishing our claim.

We first show that

$$\sum_{i \in \bar{\mathcal{I}}} x_i^\epsilon \geq \sum_{i \in \bar{\mathcal{I}}} \tilde{x}_i. \tag{60}$$

Assume to arrive at a contradiction that

$$\sum_{i \in \bar{\mathcal{I}}} x_i^\epsilon < \sum_{i \in \bar{\mathcal{I}}} \tilde{x}_i. \tag{61}$$

This implies that there exists some $j \in \bar{\mathcal{I}}$ such that $x_j^\epsilon < \tilde{x}_j$ (which also implies that $x_j^\epsilon < b_{C_j}$). We use the WE optimality conditions [Eq. (2)] for $\tilde{x}$ and x^ϵ to obtain the following:

33

- There exists some $\tilde{\lambda} \geq 0$ such that for some $i \in \bar{\mathcal{I}}$,

$$R - l_i(\tilde{x}_i) - p_i = 0 \geq \tilde{\lambda},$$

where we used the facts that $l_i(\tilde{x}_i) = 0$, $p_i = R$ [cf. Lemma 2] and $\tilde{x}_i > 0$ for some $i \in \bar{\mathcal{I}}$ [cf. Eq. (56)]. Since $\tilde{\lambda} = 0$, we have, for all $i \notin \bar{\mathcal{I}}$,

$$
\begin{aligned}
R - l_i(\tilde{x}_i) - p_i &\leq 0 \quad \text{if } \tilde{x}_i \leq b_{C_i}, \\
&\geq 0 \quad \text{if } \tilde{x}_i = b_{C_i}.
\end{aligned}
\tag{62}
$$

- There exists some $\lambda^\epsilon \geq 0$ such that

$$\epsilon - l_j(x_j^\epsilon) \leq \lambda^\epsilon, \tag{63}$$

(since $x_j^\epsilon < \tilde{x}_j$ and $p_j = R - \epsilon$), and for all $i \notin \bar{\mathcal{I}}$,

$$
\begin{aligned}
R - l_i(x_i^\epsilon) - p_i &\leq \lambda^\epsilon \quad \text{if } x_i^\epsilon = 0, \\
&\geq \lambda^\epsilon \quad \text{if } x_i^\epsilon > 0.
\end{aligned}
\tag{64}
$$

If $\lambda^\epsilon = 0$, then by Eq. (63) and the fact that $l_j(\tilde{x}_j) = 0$ (Lemma 2), we obtain

$$l_j(x_j^\epsilon) \geq \epsilon > 0 = l_j(\tilde{x}_j),$$

which is a contradiction. If $\lambda^\epsilon > 0$, then $\sum_{i=1}^{I} x_i^\epsilon = d$. Assume first that $x_i^\epsilon \leq \tilde{x}_i$ for all $i \notin \bar{\mathcal{I}}$. Then

$$\sum_{i \in \bar{\mathcal{I}}} x_i^\epsilon = d - \sum_{i \notin \bar{\mathcal{I}}} x_i^\epsilon \geq d - \sum_{i \notin \bar{\mathcal{I}}} \tilde{x}_i \geq \sum_{i \in \bar{\mathcal{I}}} \tilde{x}_i,$$

which yields a contradiction by Eq. (61). Assume next that $x_k^\epsilon > \tilde{x}_k$ for some $k \notin \bar{\mathcal{I}}$. By Eqs. (62) and (64), we have

$$R - l_k(x_k^\epsilon) - p_k \geq \lambda^\epsilon,$$
$$R - l_k(\tilde{x}_k) - p_k \leq 0,$$

which together implies that $l_k(\tilde{x}_k) > l_k(x_k^\epsilon)$, yielding a contradiction and proving Eq. (60).

Since $W(p)$ is an upper semicontinuous correspondence and the ith component of $W(p)$ is uniquely defined for all $i \notin \bar{\mathcal{I}}$, it follows that $x_i(\cdot)$ is continuous at p for all $i \notin \bar{\mathcal{I}}$. Together with Eq. (5), this implies that

$$
\begin{aligned}
\Pi(p^\epsilon, x^\epsilon) &= \sum_{i \notin \bar{\mathcal{I}}} p_i \tilde{x}_i + \sum_{i \notin \bar{\mathcal{I}}} p_i(x_i^\epsilon - \tilde{x}_i) + \sum_{i \in \bar{\mathcal{I}}} (R - \epsilon) x_i^\epsilon \\
&\geq \sum_{i \notin \bar{\mathcal{I}}} p_i \tilde{x}_i + \sum_{i \in \bar{\mathcal{I}}} (R - \epsilon) \tilde{x}_i + \sum_{i \notin \bar{\mathcal{I}}} p_i(x_i^\epsilon - \tilde{x}_i) \\
&= \sum_{i=1}^{I} p_i \tilde{x}_i - \epsilon \sum_{i \in \bar{\mathcal{I}}} \tilde{x}_i + \sum_{i \notin \bar{\mathcal{I}}} p_i(x_i^\epsilon - \tilde{x}_i) \\
&\geq \sum_{i=1}^{I} p_i \tilde{x}_i - \delta,
\end{aligned}
$$

where the last inequality holds for sufficiently small ϵ, establishing (58), and completing the proof. **Q.E.D.**

8   Appendix B: Proof of Proposition 8

We will prove Proposition 8 using Theorem 5* of Dasgupta and Maskin [10]. We start by stating a slightly simplified version of this theorem. Consider an S player game. Let the strategy space of player s, denoted by P_s, be a closed interval of $\mathbb{R}^{n_s}$ for some $n_s \in \mathbb{N}$, and its payoff function by $\pi_s(p_s, p_{-s})$. We also denote $p = (p_s, p_{-s})$, $P = \prod_{s=1}^{S} P_s$, and $P_{-s} = \prod_{k=1, k\neq s}^{S} P_s$. To state Theorem 5* in [10], we need the following three definitions.

Definition A1 Let $\pi(p) = \sum_{s\in S} \pi_s(p_s, p_{-s})$. $\pi(p)$ is upper semicontinuous in p if for all $\bar{p}$,

$$\limsup_{p\to\bar{p}} \pi(p) \leq \pi(\bar{p}).$$

Definition A2 The profit function $\pi_s(p_s, p_{-s})$ is weakly lower semicontinuous in p_s if for all $\bar{p}_s \in P_s$, there exists $\lambda \in [0, 1]$ such that for all $p_{-s} \in P_{-s}$,

$$\lambda \liminf_{p_s\downarrow\bar{p}_s} \pi_s(p_s, p_{-s}) + (1 - \lambda) \liminf_{p_s\uparrow\bar{p}_s} \pi_s(p_s, p_{-s}) \geq \pi_s(\bar{p}_s, p_{-s}).$$

Definition A3 For each player s, let $D_s \in \mathbb{N}$. For each D with $0 \leq D \leq D_s$ and each $k \neq s$ with $1 \leq k \leq S$, let f_{sk}^D be a one-to-one, continuous function. Let $\bar{P}(s)$ be a subset of P, such that

$$\bar{P}(s) = \left\{ (p_1, ..., p_S) \in P \mid \exists\ k \neq s,\ \exists\ D,\ 0 \leq D \leq D_s\ \ s.t.\ \ p_k = f_{sk}^D(p_s) \right\}.$$

In other words, $\bar{P}(s)$ is a lower dimensional subset of P (which is also of Lebesgue measure zero). Theorem 5* in [10] states:

Theorem A1 (Dasgupta-Maskin) Assume that $\pi_s(p_s, p_{-s})$ is continuous in p except on a subset P^{**} of $\bar{P}(s)$, weakly lower semicontinuous in p_s for all s and bounded, and that $\pi(p)$ is upper semicontinuous in p. Then the game $[(P_s, \pi_s); s = 1, 2, \cdots, S]$ has a mixed strategy equilibrium.

We show that our game satisfies the hypotheses of Theorem A1. We will select a function $x^*(p_s, p_{-s})$ from the set of Wardrop equilibria, $W(p_s, p_{-s})$, such that

$$\pi_s(p_s, p_{-s}) = \Pi_s(p_s, p_{-s}, x^*(p_s, p_{-s})), \qquad \forall\ s \in \mathcal{S},$$

that will satisfy these hypotheses. First, since $P_s = [0, R]^{I_s}$ and $\sum_i x_i(p) \leq d$ for all p, and all $x \in W(p)$, $\pi_s(p_s, p_{-s})$ is clearly bounded.

Since $W(p_s, p_{-s})$ is an upper semicontinuous correspondence, we select $x^*(\cdot)$ such that

$$\liminf_{p_s\uparrow\bar{p}_s} \sum_{j\in\mathcal{I}_s} x_j^*(p_s, p_{-s}) = \sum_{j\in\mathcal{I}_s} x_j^*(\bar{p}_s, p_{-s}), \qquad \forall\ \bar{p}_s \geq 0,\ \forall\ p_{-s} \geq 0. \tag{65}$$

Given $p \geq 0$, since $p_j = p_k$ for all j, $k \in \bar{\mathcal{I}}$, where $\bar{\mathcal{I}}$ is defined in Eq. (5) in Lemma 2, it follows that

$$\liminf_{p_s \uparrow \bar{p}_s} \pi_s(p_s, p_{-s}) = \pi_s(\bar{p}_s, p_{-s}), \qquad \forall \ \bar{p}_s \geq 0, \ \forall \ p_{-s} \geq 0,$$

hence ensuring that $\pi_s(p_s, p_{-s}) = \Pi_s(p_s, p_{-s}, x^*(p_s, p_{-s}))$ is weakly lower semicontinuous. We claim that we have

$$\sum_j x_j^*(p) \geq \sum_j x_j(p), \qquad \forall \ p \geq 0, \ \forall \ x \in W(p). \tag{66}$$

Assume the contrary. This implies that there exist some $\bar{p} \geq 0$, $s \in \mathcal{S}$, and $x \in W(\bar{p})$ such that

$$\sum_{j \in \mathcal{I}_s} x_j(\bar{p}) > \sum_{j \in \mathcal{I}_s} x_j^*(\bar{p}). \tag{67}$$

By Eq. (65), we have that $\sum_{j \in \mathcal{I}_s} x_j^*(p_s^n, \bar{p}_{-s}) \to \sum_{j \in \mathcal{I}_s} x_j^*(\bar{p}_s, \bar{p}_{-s})$ for some sequence $\{p_s^n\} \uparrow \bar{p}_s$. Combined with Eq. (67), this implies that

$$\sum_{j \in \mathcal{I}_s} x_j(\bar{p}) > \sum_{j \in \mathcal{I}_s} x_j^*(\hat{p}_s, \bar{p}_{-s}),$$

for some $\hat{p}_s < p_s$, contradicting the monotonicity of WE by Proposition 2.

Next, we show that $\pi_s(p_s, p_{-s})$ is continuous in p except on a set P^{**}. We define the set

$$P^{**} = \{p \mid W(p) \text{ is not a singleton}\}.$$

By the upper semicontinuity of $W(p)$, we see that $\pi_s(p_s, p_{-s})$ is continuous at all $p \notin P^{**}$. Moreover, by Lemma 2, it follows that $P^{**} \subset \bar{P}$, where

$$\bar{P} = \{p \mid p_j = p_k, \text{ for some } j \neq k\} \cup \{p \mid p_j = R, \text{ for some } j\},$$

which is a lower dimensional set. This establishes the desired condition for Theorem A1.

Finally, we show that

$$\pi(p) = \sum_{s \in \mathcal{S}} \pi_s(p_s, p_{-s}) = \sum_{i \in \mathcal{I}} p_i x_i^*(p),$$

is continuous at all p. Given some $p \geq 0$, define $\bar{\mathcal{I}}$ as in Eq. (5) of Lemma 2. If $\bar{\mathcal{I}} = \emptyset$, then we automatically have that π is continuous at p. Assume that $\bar{\mathcal{I}} \neq \emptyset$. Since $x_i^{OE}(\cdot)$ is continuous at p for all $i \notin \bar{\mathcal{I}}$ and $p_j = p_k$ for all j, $k \in \bar{\mathcal{I}}$, it is sufficient to show that $\sum_{i \in \bar{\mathcal{I}}} x_i^*(p)$ is continuous at p, i.e., for a sequence $\{p^n\}$ with $p^n \in [0, \mathbb{R}]^I$ and $p^n \to p$, we show that

$$\lim_{n \to \infty} \sum_{i \in \bar{\mathcal{I}}} x_i^*(p^n) = \sum_{i \in \bar{\mathcal{I}}} x_i^*(p).$$

Define

$$\tilde{d}(p^n) = \sum_{i \notin \bar{\mathcal{I}}} x_i^*(p^n).$$

Since $x_i(\cdot)$ is continuous at p for all $i \notin \bar{\mathcal{I}}$, we have $\tilde{d}(p^n) \to d(p) = \sum_{i \notin \bar{\mathcal{I}}} x_i(p)$. Consider two cases:

36

- $\sum_{i \in \tilde{\mathcal{I}}} b_{C_i} > d - \tilde{d}(p)$. Since $x^*(p)$ is the maximum l_1-norm element of $W(p)$ [cf. Eq. (66)] and $l_i(x_i^*) = 0$ for all $i \in \mathcal{I}$, this implies that $\sum_{i \in \mathcal{I}} x_i^*(p) = d$ and for all n sufficiently large $\sum_{i \in \mathcal{I}} x_i^*(p^n) = d$, establishing the claim.

- $\sum_{i \in \tilde{\mathcal{I}}} b_{C_i} \leq d - \tilde{d}(p)$. By the same reasoning as in the previous part, this implies that $\sum_{i \in \mathcal{I}} x_i^*(p) = \sum_{i \in \tilde{\mathcal{I}}} b_{C_i}$. Moreover, for all $\epsilon > 0$, there exists some n sufficiently large such that

$$\left| \sum_{i \in \mathcal{I}} x_i^*(p^n) - \sum_{i \in \tilde{\mathcal{I}}} b_{C_i} \right| \leq \epsilon,$$

establishing the claim.

The preceding enable us to apply the theorem, completing the proof. **Q.E.D.**

9 Appendix C: Proof of Lemma 6

We first prove the following lemma:

Lemma 10 Let (p^{OE}, x^{OE}) be an OE such that $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$. Let Assumptions 1 and 2 hold. If $p_j^{OE} x_j^{OE} > 0$ for some $j \in \mathcal{I}$, then $W(p^{OE})$ is a singleton.

Proof. Since $p_j^{OE} x_j^{OE} > 0$ for some $j \in \mathcal{I}$, it follows, by Lemma 5, that $p_i^{OE} x_i^{OE} > 0$ for all $i \in \mathcal{I}$. We first show that for all $x \in W(p^{OE})$, we have $x_i \leq x_i^{OE}$ for all i. If $l_i(x_i^{OE}) > 0$, then by Lemma 2, $x_i = x_i^{OE}$ for all $x \in W(p^{OE})$. If $l_i(x_i^{OE}) = 0$, then $\mathcal{I}_s = \{i\}$ for some s by the fact that $x_i^{OE} > 0$ and Assumption 2, which implies that $x_i \leq x_i^{OE}$ by the definition of an OE (cf. Definition 10).

Since $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$, we have $\sum_{i=1}^{I} x_i^{OE} = d$. Moreover, the fact that $x_i \leq x_i^{OE}$ for all $x \in W(p^{OE})$ implies that $\min_j \{p_j^{OE} + l_j(x_j)\} < R$ as well, and therefore $\sum_{i=1}^{I} x_i = d$, showing that $x_i = x_i^{OE}$ for all $x \in W(p^{OE})$, for all $i \in \mathcal{I}$. **Q.E.D.**

Proof of Lemma 6. We first prove this result for a network with two links. Assume to arrive at a contradiction that

$$l_2^+(x_2^{OE}) > l_2^-(x_2^{OE}). \tag{68}$$

Let $\{\epsilon^k\}$ be a scalar sequence with $\epsilon^k \downarrow 0$. Consider the sequence $\{x_1(\epsilon^k)\}$ where $x_1(\epsilon^k)$ is the load of link 1 at a WE given price vector $(p_1^{OE} + \epsilon^k, p_2^{OE})$. By Proposition 1 and Lemma 10, the WE correspondence $W(p)$ is upper-semicontinuous and $W(p^{OE})$ is a singleton. Therefore, it follows that $x_1(\epsilon^k) \to x_1^{OE}$. Define

$$\frac{\partial^+ x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} = \lim_{k \to \infty} \frac{x_1(\epsilon^k) - x_1^{OE}}{\epsilon^k}. \tag{69}$$

Similarly, let $x_1(-\epsilon^k)$ be the load of link 1 at a WE given price vector $(p_1^{OE} - \epsilon^k, p_2^{OE})$. Since $W(p^{OE})$ is a singleton, we also have $x_1(-\epsilon^k) \to x_1^{OE}$. Define

$$\frac{\partial^- x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} = \lim_{k \to \infty} \frac{x_1^{OE} - x_1(-\epsilon^k)}{\epsilon^k}. \tag{70}$$

Since $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$, it can be seen using Lemma 4 that

$$\frac{\partial^+ x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} \geq \frac{-1}{l_1^-(x_1^{OE}) + l_2^+(x_2^{OE})},$$

and

$$\frac{\partial^- x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} \leq \frac{-1}{l_1^+(x_1^{OE}) + l_2^-(x_2^{OE})}.$$

Since $l_1^+(x_1^{OE}) = l_1^-(x_1^{OE})$ by Assumption 3, this combined with Eq. (68) yields

$$\frac{\partial^+ x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} > \frac{\partial^- x_1(p_1^{OE}, p_2^{OE})}{\partial p_1}. \tag{71}$$

Consider the profit of service provider 1, $\Pi_1(p_1^{OE}, p_2^{OE}) = p_1^{OE} x^{OE}$. Define

$$\frac{\partial^+ \Pi_1(p_1^{OE}, p_2^{OE})}{\partial p_1} = \lim_{k \to \infty} \frac{\Pi_1(p_1^{OE} + \epsilon^k, p_2^{OE}) - \Pi_1(p_1^{OE}, p_2^{OE})}{\epsilon^k},$$

$$\frac{\partial^- \Pi_1(p_1^{OE}, p_2^{OE})}{\partial p_1} = \lim_{k \to \infty} \frac{\Pi_1(p_1^{OE}, p_2^{OE}) - \Pi_1(p_1^{OE} - \epsilon^k, p_2^{OE})}{\epsilon^k}.$$

Since p_1^{OE} is a maximum of $\Pi_1(\cdot, p_2^{OE})$, we have

$$\frac{\partial^+ \Pi_1(p_1^{OE}, p_2^{OE})}{\partial p_1} = x_1^{OE} + p_1^{OE} \frac{\partial^+ x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} \leq 0, \tag{72}$$

and

$$\frac{\partial^- \Pi_1(p_1^{OE}, p_2^{OE})}{\partial p_1} = x_1^{OE} + p_1^{OE} \frac{\partial^- x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} \geq 0, \tag{73}$$

which, when combined, yields

$$\frac{\partial^+ x_1(p_1^{OE}, p_2^{OE})}{\partial p_1} \leq \frac{\partial^- x_1(p_1^{OE}, p_2^{OE})}{\partial p_1}, \tag{74}$$

which is a contradiction by Eq. (71), thus showing that we have $l_2^+(x_2^{OE}) = l_2^-(x_2^{OE})$.

We next consider a network with multiple links. As in Eqs. (69) and (70), we define for all $i \in \mathcal{I}$,

$$\frac{\partial^+ x_i(p^{OE})}{\partial p_1} = \lim_{k \to \infty} \frac{x_i(\epsilon^k) - x_i^{OE}}{\epsilon^k},$$

$$\frac{\partial^- x_i(p^{OE})}{\partial p_1} = \lim_{k \to \infty} \frac{x_i^{OE} - x_i(-\epsilon^k)}{\epsilon^k}.$$

Using the same line of argument as above, we obtain

$$\frac{\partial^+ x_1(p^{OE})}{\partial p_1} \geq \frac{-1}{l_1^-(x_1^{OE}) + \frac{1}{\sum_{i \neq 1} \frac{1}{l_i^+(x_i^{OE})}}},$$

$$\frac{\partial^- x_1(p^{OE})}{\partial p_1} \leq \frac{-1}{l_1^+(x_1^{OE}) + \frac{1}{\sum_{i \neq 1} \frac{1}{l_i^-(x_i^{OE})}}}. \tag{75}$$

Let $1 \in \mathcal{I}_s$, and without loss of any generality, assume that all l_i's for $i \in \mathcal{I}_s$ are smooth (recall Assumption 3). For all $i \in \mathcal{I}_s$, $i \neq 1$, we obtain

$$\frac{\partial^+ x_i(p^{OE})}{\partial p_1} \geq \frac{1}{l_i^+(x_i^{OE}) \left(1 + l_1^-(x_1^{OE}) \left(\sum_{j \neq 1} \frac{1}{l_j^+(x_j^{OE})}\right)\right)},$$

$$\frac{\partial^- x_i(p^{OE})}{\partial p_1} \leq \frac{1}{l_i^-(x_i^{OE}) \left(1 + l_1^+(x_1^{OE}) \left(\sum_{j \neq 1} \frac{1}{l_j^-(x_j^{OE})}\right)\right)}.$$

To arrive at a contradiction, assume that $l_j^+(x_j^{OE}) > l_j^-(x_j^{OE})$ for some $j \notin \mathcal{I}_s$. Then the preceding two sets of equations imply that

$$\frac{\partial^+ x_i(p^{OE})}{\partial p_1} > \frac{\partial^- x_i(p^{OE})}{\partial p_1}. \tag{76}$$

for all $i \in \mathcal{I}_s$.

Next, Eqs. (72) and (73) for multiple link case are given by

$$\frac{\partial^+ \Pi_1(p^{OE})}{\partial p_1} = x_1^{OE} + p_1^{OE} \frac{\partial^+ x_1(p^{OE})}{\partial p_1} + \sum_{i \in \mathcal{I}_s, i \neq 1} p_i^{OE} \frac{\partial^+ x_i(p^{OE})}{\partial p_1} \leq 0,$$

$$\frac{\partial^- \Pi_1(p^{OE})}{\partial p_1} = x_1^{OE} + p_1^{OE} \frac{\partial^- x_1(p^{OE})}{\partial p_1} + \sum_{i \in \mathcal{I}_s, i \neq 1} p_i^{OE} \frac{\partial^- x_i(p^{OE})}{\partial p_1} \geq 0, \tag{77}$$

which are inconsistent with Eq. (76), leading to a contradiction. This proves the claim for the multiple link case. **Q.E.D.**

10   Appendix D: Proof of Proposition 9

We first assume that $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} < R$. Consider service provider s and assume without loss of generality that $1 \in \mathcal{I}_s$. Since $p_j^{OE} x_j^{OE} > 0$ for some $j \in \mathcal{I}_{s'}$ and $s' \in \mathcal{S}$, it follows by Lemma 5 that $p_i^{OE} x_i^{OE} > 0$ for all $i \in \mathcal{I}$. Together with Lemma 4, this implies that $((p_i^{OE})_{i \in \mathcal{I}_s}, x^{OE})$ is an optimal solution of the problem

$$\text{maximize}_{((p_i)_{i \in \mathcal{I}_s}, x) \geq 0} \quad \sum_{i \in \mathcal{I}_s} p_i x_i \tag{78}$$

$$\text{subject to} \quad l_1(x_1) + p_1 = l_i(x_i) + p_i, \qquad i \in \mathcal{I}_s - \{1\},$$

$$l_1(x_1) + p_1 = l_i(x_i) + p_i^{OE}, \qquad i \notin \mathcal{I}_s,$$

$$l_1(x_1) + p_1 \leq R, \tag{79}$$

$$\sum_{i \in \mathcal{I}} x_i \leq d.$$

39

By Lemma 6, we have that l_i is continuously differentiable in a neighborhood of x_i^{OE} for all i (since the gradient mapping of a convex function is continuous over the set the function is differentiable, see Rockafellar [29]). Therefore, by examining the Karush-Kuhn-Tucker conditions of this problem, we obtain

$$p_i^{OE} = x_i^{OE} l_i'(x_i^{OE}) - \theta, \qquad \forall\, i \in \mathcal{I}_s, \tag{80}$$

where

$$\theta = \begin{cases} 0, & \text{if } l_j'(x_j^{OE}) = 0 \text{ for some } j \notin \mathcal{I}_s, \\ -\dfrac{\sum_{j \in \mathcal{I}_s} x_j^{OE}}{\sum_{j \notin \mathcal{I}_s} \frac{1}{l_j'(x_j^{OE})}}, & \text{otherwise,} \end{cases} \tag{81}$$

showing the result in Eq. (16).

We next assume that $\min_j \{p_j^{OE} + l_j(x_j^{OE})\} = R$. Using the assumption that $p_j^{OE} x_j^{OE} > 0$ for some $j \in \mathcal{I}$ and Lemma 4, this implies that

$$p_i^{OE} = R - l_i(x_i^{OE}), \qquad \forall\, i,$$

and thus for all $s \in \mathcal{S}$, x^{OE} is an optimal solution of

$$
\begin{aligned}
\text{maximize} \quad & \sum_{i \in \mathcal{I}_s} (R - l_i(x_i)) x_i \\
\text{subject to} \quad & x_i \in T_i, \qquad \forall\, i \notin \mathcal{I}_s \\
& \sum_{i \in \mathcal{I}} x_i \leq d,
\end{aligned}
\tag{82}
$$

where $T_i = \{x_i \mid p_i^{OE} + l_i(x_i) = R\}$ is either a singleton or a closed interval. Since this is a convex problem, using the optimality conditions, we obtain

$$R - l_i(x_i^{OE}) - x_i^{OE} g_{l_i} = \theta_s, \qquad \forall\, i \in \mathcal{I}_s,$$

where $\theta_s \geq 0$ is the Lagrange multiplier associated with constraint (82), and $g_{l_i} \in \partial l_i(x_i^{OE})$. Since $l_i^-(x_i^{OE}) \leq g_{l_i}$, the preceding implies

$$p_i^{OE} = R - l_i(x_i^{OE}) \geq x_i^{OE} l_i^-(x_i^{OE}), \qquad \forall\, i \in \mathcal{I},$$

proving (17).

To prove (18), consider some $i \in \mathcal{I}$ with $\mathcal{I}_s = \{i\}$ for some s and the sequence of price vectors $\{p^k\}$ with $p^k = (p_i^{OE} - \epsilon^k, p_{-i}^{OE})$. Let $\{x^k\}$ be a sequence such that $x^k \in W(p^k)$ for all k. By the upper semicontinuity of $W(p)$, it follows that $x^k \to \bar{x}$ with $\bar{x} \in W(p^{OE})$ and $\bar{x} \leq x^{OE}$ (see the proof of Lemma 10). Moreover, by Lemma 2, we have $x_i^k \geq x_i^{OE}$ for all k, which implies that $\bar{x}_i \geq x_i^{OE}$, showing that $x_i^k \to x_i^{OE}$. We can now use Eqs. (73) and (77) (by substituting i instead of 1 and using $\mathcal{I}_s = \{i\}$) to conclude that

$$p_i^{OE} \leq x_i^{OE} l_i^+(x_i^{OE}) + \frac{x_i^{OE}}{\sum_{j \neq 1} \frac{1}{l_j^-(x_j^{OE})}}.$$

Q.E.D.

References

[1] Acemoglu, D., and Ozdaglar, A.,"Flow control, routing, and performance from service provider viewpoint," *LIDS report*, WP-1696, May 2004.

[2] Anshelevich E., Dasgupta A., Kleinberg J., Tardos E., Wexler T., and Roughgarden T., "The price of stability for network design with selfish agents," *IEEE Symposium on Foundations of Computer Science,* pp. 295-304, 2004.

[3] Basar, T. and Srikant R., "Revenue-maximizing pricing and capacity expansion in a many-users regime," *Proceedings of INFOCOM*, 2002.

[4] Beckmann, M., Mcguire, C. B., and Winsten, C. B., Studies in the Economics of Transportation. Yale University Press, 1956.

[5] Berge, C., Topological Spaces, Dover reprint, 1963.

[6] Benassy, J. P., "Market size and substitutability in imperfect competition: A Bertrand-Edgeworth- Chamberlin model," *Review of Economic Studies*,vol. 56, pp. 217-234, 1989.

[7] Bertsekas, D. P., Nedic, A., and Ozdaglar, A., Convex Analysis and Optimization, Athena Scientific, Belmont, MA, 2003.

[8] Correa, J. R., Schulz, A. S., and Stier Moses, N., "Selfish Routing in Capacitated Networks," *Mathematics of Operations Research*, 29:4, pp. 961-976, Nov. 2004.

[9] Dafermos, S. and Sparrow F. T., "The traffic assignment problem for a general network," *Journal of Research of the National Bureau of Standards-B. Mathematical Sciences,* vol. 73(2), pp. 91-118, 1969.

[10] Dasgupta, P. and Maskin E., "The existence of equilibrium in discontinuous economic games. 2: Theory," *Review of Economic Studies*, vol. 53, pp. 1-26, 1986.

[11] Edgeworth, F., "The pure theory of monopoly," in Papers relating to Political Economy, vol. 1, pp. 111-143, London, Mc Millan, 1925.

[12] Friedman, E., "A Generic Analysis of Selfish Routing," *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004.

[13] Fudenberg, D. and Tirole J., Game Theory. The MIT Press, 1991.

[14] Hart, O. D., "Monopolistic competition in a large economy with differentiated commodities," *Review of Economic Studies*, vol. 46, pp. 1-30, 1979.

[15] He, L. and Walrand, J., "Pricing internet services with multiple providers," *Proceedings of Allerton Conference*, 2003.

[16] Johari, R. and Tsitsiklis, J., "Network resource allocation and a congestion game," *Mathematics of Operations Research*, to appear, 2004.

[17] Kelly, F. P., Maulloo A. K., and Tan D. K.,"Rate control for communication networks: shadow prices, proportional fairness, and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237-252, 1998.

[18] Korilis, Y. A., Lazar, A. A., and Orda, A.,"Achieving network optima using Stackelberg routing strategies," *IEEE/ACM Transactions on Networking*, vol. 5(1), pp. 161-173, 1997.

[19] Koutsoupias, E. and Papadimitriou, C.,"Worst-case Equilibria," *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science,* pp. 404-413, 1999.

[20] Low, S. and Lapsley, D. E., "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7(6), pp. 861-874, 1999.

[21] Mascolell, A., Whinston, M. D., and Green, J. R., Microeconomic Theory, Oxford, United Kingdom: Oxford University Press, 1995.

[22] Novshek, W., "Perfectly competitive markets as the limits of Cournot markets," *Journal of Economic Theory,* vol. 19, pp. 223-266, 1985.

[23] Orda, A., Rom, R., and Shimkin, N., "Competitive routing in multi-user communication networks," *IEEE/ACM Transactions on Networking*, vol. 1, pp. 510-521, 1993.

[24] Osborne and Rubinstein, Game Theory, The MIT Press.

[25] Patriksson, M., The Traffic Assignment Problem: Models and Methods, VSP BV, Netherlands, 1994.

[26] Perakis, G., "The Price of Anarchy when Costs are Non-separable and Asymmetric", manuscript.

[27] Pigou, A. C., *The Economics of Welfare.* Macmillan, 1920.

[28] Roberts, D. and Postlewaite, A., "The incentives for price taking in large exchange economies," Econometrica, vol. 44, pp. 115-127, 1976.

[29] Rockafeller, Convex Analysis, Princeton, New Jersey: Princeton University Press, 1970.

[30] Rosen, J. B., "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica* vol. 33(3), pp. 520-534, 1965.

[31] Roughgarden, T. and Tardos, E., "How bad is selfish routing?" *Journal of the ACM*, vol. 49(2), pp. 236-259, 2002.

[32] Sanghavi S. and Hajek B., "Optimal allocation of a divisible good to strategic buyers," 43rd IEEE Conference on Decision and Control - CDC 2004.

[33] Samuelson, P. A., "Spatial price equilibrium and linear programming," *American Economic Review*, vol. 42, pp. 283-303, 1952.

[34] Shenker, S., "Fundamental design issues for the future Internet," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1176-1188, 1995.

[35] Shubik, M., Strategy and Market Structure, New York: Wiley, 1959.

[36] Tirole, J., The Theory of Industrial Organization, Cambridge: MIT Press, MA, 1990.

[37] Topkis, D., "Equilibrium points in nonzero-sum n-person submodular games," *SIAM Journal of Control and Optimization,* vol. 17, pp. 773-787, 1979.

[38] Veblen, T., The Theory of the Leisure Class: An Economic Study of Institutions, New York: Vanguard Press.

[39] Vives, X., "On the efficiency of Bertrand and Cournot equilibria with product differentiation," *Journal of Economic Theory,* vol. 36, pp. 166-175, 1985.

[40] Vives, X., Oligopoly Pricing, Cambridge: MIT Press, MA, 2001.

[41] Wardrop, J. G., "Some theoretical aspects of road traffic research," *Proceedings of the Institute of Civil Engineers*, Pt. II, vol. 1, pp. 325-378, 1952.

[42] Wilson, R. B., Nonlinear Pricing, New York: Oxford University Press, 1993.